Dynamics of state price densities

Härdle, Wolfgang; Hlávka, Zdeněk

Dynamics of state price densities

Wolfgang Härdle, Zdeněk Hlávka

PII: S0304-4076(09)00024-4
DOI: 10.1016/j.jeconom.2009.01.005
Reference: ECONOM 3158

To appear in: *Journal of Econometrics*

Received date: 11 January 2009
Accepted date: 12 January 2009

Dynamics of State Price Densities

Wolfgang Härdle [a], Zdeněk Hlávka [b,*]

[a] *CASE—Center for Applied Statistics and Economics*
Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin
Spandauer Str. 1, 10178 Berlin, Germany

[b] *Charles University in Prague, Department of Statistics*
Sokolovská 83, 18675 Praha, Czech Republic

Abstract

State price densities (SPDs) are an important element in applied quantitative finance. In a Black-Scholes world they are lognormal distributions but in practice volatility changes and the distribution deviates from log-normality. In order to study the degree of this deviation, we estimate SPDs using EUREX option data on the DAX index via a nonparametric estimator of the second derivative of the (European) call pricing function. The estimator is constrained so as to satisfy no-arbitrage constraints and corrects for the intraday covariance structure in option prices. In contrast to existing methods, we do not use any parametric or smoothness assumptions.

Key words: Option pricing, State price density, Nonlinear least squares, Constrained estimation

JEL classification: C13, C14, G13

1 Introduction

The dynamics of option prices carries information on changes in state price densities (SPDs). The SPD contains important information on the behavior and expectations of the market and is used for pricing and hedging. The most important application of SPD is that it allows to price options with

* Corresponding author. Tel.:+420 221 913 284, fax:+420 283 073 341
Email addresses: haerdle@wiwi.hu-berlin.de (Wolfgang Härdle), hlavka@karlin.mff.cuni.cz (Zdeněk Hlávka).

ACCEPTED MANUSCRIPT

complicated payoff functions simply by (numerical) integration of the payoff with respect to this density.

Prices $C_t(K,T)$ of European options with strike price K observed at time t and expiring at time T allow to deduce the state price density $f(.)$ using the relationship (Breeden and Litzenberger, 1978):

$$f(K) = \exp\{r(T-t)\}\frac{\partial^2 C_t(K,T)}{\partial K^2}. \tag{1}$$

Equation (1) can be used to estimate SPD $f(K)$ from the observed option prices. An extensive overview of parametric and other estimation techniques can be found, e.g., in Jackwerth (1999). An application to option pricing is given in Buehler (2006).

Kernel smoothers were in this framework proposed and successfully applied by, e.g., Aït-Sahalia and Lo (1998), Aït-Sahalia and Lo (2000), Aït-Sahalia, Wang, and Yared (2000) or Huynh, Kervella, and Zheng (2002). Aït-Sahalia and Duarte (2003) proposed a method for nonparametric estimation of the SPD under constraints like positivity, convexity, and boundedness of the first derivative. Bondarenko (2003) calculates arbitrage-free SPD estimates using positive convolution approximation (PCA) methodology and demonstrates its properties in a Monte Carlo studied based on closing prices of the S&P 500 options. Another sophisticated approach based on smoothing splines allowing to include these constraints is described and applied on simulated data in Yatchew and Härdle (2006). In the majority of these papers, the focus was more on the smoothing techniques rather than on a no-arbitrage argument although a crucial element of local volatility models is the absence of arbitrage (Dupire, 1994). Highly numerically efficient pricing algorithms, e.g., by Andersen and Brotherton-Ratcliffe (1997), rely heavily on no-arbitrage properties. Kahalé (2004) proposed a procedure that requires solving a set of nonlinear equations with no guarantee of a unique solution. Moreover, for that algorithm the data feed already is (unrealistically) expected to be arbitrage free (Fengler, 2005; Fengler, Härdle, and Mammen, 2007). In addition, the covariance structure of the quoted option prices (Renault, 1997) is rarely incorporated into the estimation procedure.

Insert Table 1

In Table 1, we give an overview of selected properties of different estimation techniques. The parametric approach may be used to estimate parameters of a probability density lying in some preselected family. The parametric models may be further extended by considering more flexible probability densities or mixtures of distributions. Approaches based on nonparametric smoothing techniques are more flexible since the shape of a nonparametric SPD estimate is not fixed in advance and the method controls only the smoothness of the estimate. For example, the smoothness of a kernel regression estimator depends mostly on the choice of the bandwidth parameter, the smoothness of the PCA estimator (Bondarenko, 2003) depends on the choice of the kernel, the smoothness of the NNLS estimator (Yatchew and Härdle, 2006) is controlled by constraining the Sobolev norm of the SPD; using these nonparametric estimators, systematic bias may typically occur in case of oversmoothing. Constraints on estimators are more easily implemented for globally valid parametric models than for local (nonparametric) models. The use of a standard smoothing technique which does not account for the constraints is not advisable. The value of the nonparametric estimate cannot be calculated in regions without any data and, therefore, the support of the nonparametrically estimated SPDs is limited by the range of the observed strike prices even for the nonparametric-under-constraints techniques.

Most of the commonly used estimation techniques do not specify explicitely the source of random error in the observed option prices, see Renault (1997) for an extensive review of this subject. A common approach in the SPD estimation is to use either the closing option prices or to correct the intraday option prices by the current value of the underlying asset. Both approaches lack interpretation if the shape of the SPD changes rapidly. This can be made clear by a gedankenexperiment: if the shape of the SPD changes dramatically during the day, correcting the observed option prices by the value of the underlying asset and then estimating the SPD would lead to an estimate of some (nonexisting) daily average of the true SPDs. We try to circumvent this problem by introducing a simple model for the intraday covariance structure of option prices which allows us to estimate the value of the true SPD at an arbitrarily chosen fixed time, see also Hlávka and Svojík (2008). Most often, we are interested in the estimation of the current SPD.

We develop a simple estimation technique in order to construct constrained SPD estimates from the observed intraday option prices which are treated as repeated observations collected during a certain time period. The proposed technique involves constrained LS-estimation, it enables us to construct confidence intervals for the current value of SPD and prediction intervals for its future development, and it does not depend on any tuning (smoothness) parameter. The construction of a simple approximation of the covariance structure of the observed option prices follows naturally from the derivation of our nonparametric constrained estimator. This covariance structure is interesting in itself, it separates two sources of random errors, and it is applicable to other SPD estimators.

We study the development of the estimated SPDs in Germany over 8 years. A no-arbitrage argument is imposed at each time point leading (mathematically) to the above mentioned no-arbitrage constraints. This, of course, is a vital feature for the trading purposes where the derived (implied) volatility surfaces for different strikes and maturities are needed for proper judgment of risk and return.

The resulting SPDs and implied volatility surfaces are not smooth per se. In most applications, this is not a disadvantage though since, first, we may smooth the resulting SPD estimates (Hlávka and Svojík, 2008) and, second, we are mostly interested in functionals of the estimated SPD like, e.g., the expected payoff or the forward price. Another important feature that can be easily estimated from the nonsmooth SPDs are the quantiles, see Section 6.2 for an application.

In Section 2, we introduce notation, discuss constraints that are necessary for estimating SPDs, and we construct a very simple unconstrained SPD estimator using simple linear regression. In Section 3, this estimator is modified so that it satisfies the shape constraints given in Section 2.1. We demonstrate that the covariance structure of the option prices exhibits correlations depending both on the strike price and time of the trade in Section 4. In Section 5, we apply our estimation technique on option prices observed in year 1995 and we show that the proposed approximation of the covariance structure removes dependency and heteroscedasticity of the residuals. The dynamics of the estimated SPDs in years 1995–2003 is studied in Section 6.

2 Construction of the estimate

The fair price of a European call option with payoff $(S_T - K)_+ = \max(S_T - K, 0)$, with S_T denoting the price of the stock at time T, t the current time, K the strike price, and r the risk free interest rate, can be written as:

$$C_t(K,T) = \exp\{-r(T-t)\} \int_0^\infty (S_T - K)_+ f(S_T) dS_T, \tag{2}$$

i.e., as the discounted expected value of the payoff with respect to the SPD $f(.)$. For the sake of simplicity of the following presentation, we assume in the rest of the paper that the discount factor $\exp\{-r(T-t)\} = 1$. In applications, this is achieved by correcting the observed option prices by the known risk free interest rate r and the time to maturity $(T-t)$ in (2). At the time of the trade, the current index price and volatility are common to all options and, hence, do not appear explicitly in equation (2).

Let us denote the i-th observation of the strike price by K_i and the corresponding option price, divided by the discount factor $\exp\{-r(T-t)\}$ from (2), by $C_i = C_{t,i}(K_i, T)$. In practice, on any given day t, one observes option prices repeatedly for a small number of distinct strike prices. Therefore, it is useful to adopt the following notation. Let $\mathcal{C} = (C_1, \dots, C_n)^\top$ be the vector of the observed option prices on day t sorted by strike price. Then, the vector of strike prices has the following structure:

$$\mathcal{K} = \begin{pmatrix} K_1 \\ K_2 \\ \vdots \\ K_n \end{pmatrix} = \begin{pmatrix} k_1 \mathbf{1}_{n_1} \\ k_2 \mathbf{1}_{n_2} \\ \vdots \\ k_p \mathbf{1}_{n_p} \end{pmatrix},$$

where $k_1 < k_2 < \cdots < k_p$, $n_j = \sum_{i=1}^{n} \boldsymbol{I}(K_i = k_j)$ with $\boldsymbol{I}(.)$ denoting the indicator function and $\mathbf{1}_n$ a vector of ones of length n.

2.1 Assumptions and constraints

Let us now concentrate on options corresponding to a single maturity T observed at fixed time t. Let us assume that the i-th observed option price (corresponding to strike price K_i) follows the model

$$C_{t,i}(K_i, T) = \mu(K_i) + \varepsilon_i, \tag{3}$$

where ε_i are iid random variables with zero mean and variance σ^2. In practice, one might expect that the errors exhibit correlations depending on strike price and time. Heteroscedasticity can be incorporated in model (3) if we assume that the random errors ε_i have variance $\operatorname{Var} \varepsilon_i = \sigma^2_{K_i}$ leading to weighted least squares. The assumptions on the distribution of random errors will be investigated in more detail in Subsection 5.3. Following Renault (1997), we interpret the observed option price as the price given by a pricing formula plus an error term and in Section 4, we suggest a covariance structure for the observed option prices taking into account the dependencies across strike prices and times of trade.

Harrison and Pliska (1981) characterized the absence of arbitrage by the existence of a unique risk neutral SPD $f(.)$. From formula (2) and properties of a probability density it follows that, in a continuous setting, the function $\mu(.)$, defined on $\mathbb{R}^+$, has to satisfy the following no-arbitrage constraints:

1': it is positive,
2': it is decreasing in K,
3': it is convex,
4': its second derivative exists and it is a density (i.e., nonnegative and it integrates to one).

Let us now have a look at functions satisfying Constraints 1'–4'.

LEMMA 1 *Suppose that $\mu : \mathbb{R}^+ \to \mathbb{R}^+$ satisfies Constraints 1'–4'. Then the first derivative, $\mu^{(1)}(.)$, is nondecreasing and such that $\lim_{x\to 0} \mu^{(1)}(x) = -1$ and $\lim_{x\to +\infty} \mu^{(1)}(x) = 0$.*

Proof:
Constraint 4' implies that the first derivative, $\mu^{(1)}$, exists and that it is differentiable. $\lim\limits_{x\to+\infty} \mu^{(1)}(x)$ exists since the function $\mu^{(1)}$ is nondecreasing (Constraint 3') and bounded (Constraint 2'). Next, $\lim\limits_{x\to\infty} \mu^{(1)}(x) = 0$ since a negative limit would violate Constraint 1' for large x ($\mu^{(1)}(x)$ cannot be positive since $\mu(x)$ is decreasing). Finally, Constraint 4', $1 = \int\limits_0^\infty \mu^{(2)}(x)dx = \lim\limits_{x\to+\infty} \mu^{(1)}(x) - \lim\limits_{x\to 0} \mu^{(1)}(x)$, implies that $\lim\limits_{x\to 0} \mu^{(1)}(x) = -1$. □

REMARK 1 *Lemma 1 allows us to restate Constraints 3' and 4' in terms of $\mu^{(1)}(.)$ by assuming that $\mu^{(1)}(.)$ is differentiable, nondecreasing, and such that $\lim_{x\to 0} \mu^{(1)}(x) = -1$ and $\lim_{x\to+\infty} \mu^{(1)}(x) = 0$.*

In this section, we stated only constrains guaranteeing that the SPD estimate will be a probability density. Constraints for the expected value of the SPD estimate are discussed in Subsection 3.6.

2.2 Existence and uniqueness

In this subsection we address the issue of existence and uniqueness of a regression function, $\hat{C}(.)$, satisfying the required assumptions and constraints. In practice, we don't deal with a continuous function. Hence, we restate Constraints 1'–4' for discrete functions, defined only on a finite set of distinct points, say $k_1 < \cdots < k_p$, in terms of their function values, $C(k_i)$, and their scaled first differences, $C^{(1)}_{k_i,k_j} = \{C(k_i) - C(k_j)\}/\{k_i - k_j\}$.

1: $C(k_i) \geq 0$, $i = 1, \ldots, p$,
2: $k_i < k_j$ implies that $C(k_i) \geq C(k_j)$,
3: $k_i < k_j < k_l$ implies that $-1 \leq C^{(1)}_{k_i,k_j} \leq C^{(1)}_{k_j,k_l} \leq 0$.

It is easy to see that Constraints 1–2 are discrete versions of Constraints 1' and 2'. Constraint 3 is a discrete version of Constraints 3' and 4', see Remark 1.

From now on, similarly as in Robertson, Wright, and Dykstra (1988), we think of the collection, $\mathcal{C}$, of functions satisfying Constraints 1–3 as a subset of a p-dimensional Euclidean space, where p is the number of distinct k_i's. The constrained regression, $\hat{C}$, is in this setting the closest point of $\mathcal{C}$ to the vector C of the observed option prices with distances measured by the usual

Euclidean distance

$$d(f, C) = (f - C)^\top (f - C) = \sum_{i=1}^{n} \{f(K_i) - C(K_i)\}^2. \tag{4}$$

From this point of view, the regression function, $\hat{C}$, consists only of the values of the function in the points $k_1, \ldots, k_p$. The first and second differences are used to approximate the first and the second derivatives, respectively.

We claim that the set, $\mathcal{C}$, of functions satisfying Constraints 1–3, is closed in the topology induced by the metric given by Euclidean distance and it is convex, i.e., if $f, g \in \mathcal{C}$ and $0 \le a \le 1$, then $af + (1-a)g \in \mathcal{C}$.

LEMMA 2 *If $\hat{C} \in \mathcal{C}$ is the regression of $C(K_i)$, $i = 1, \ldots, n$, on $k_1 < \cdots < k_p$ under Constraints 1–3 and if a and b are constants such that $a \le C(K_i) \le b$, $\forall i$, then $a \le \hat{C}(k_i) \le b + (k_p - k_1)$.*

Proof:
It is not possible that $\hat{C}(k_i)$ lies above b for all k_i's (otherwise we would get a better fit only by shifting $\hat{C}(k_i)$). The upper bound now follows from Constraint 3.

The validity of the lower bound may be demonstrated similarly. Clearly, it is not possible that $\hat{C}(k_i)$ lie below a for all k_i's. Moreover, it is not possible that $\hat{C}(k_1) \ge \cdots \ge \hat{C}(k_i) \ge a > \hat{C}(k_{i+1}) \ge \cdots \ge \hat{C}(k_p)$ for any i since in such situation the fit could be trivially improved by increasing $\hat{C}(k_{i+1}), \ldots, \hat{C}(k_p)$ by some small amount, e.g., by $a - \hat{C}(k_{i+1})$ without violating any of the Constraints 1–3. $\square$

THEOREM 1 *A regression, $\hat{C} = \arg\min_{f \in \mathcal{C}} d(f, C)$, satisfying Constraints 1–3, exists and it is unique.*

Proof:
Lemma 2 implies that $\hat{C}$ belongs to a subset, $\mathcal{S}$, of $\mathcal{C}$ bounded below by a and above by $b + (k_p - k_1)$. Thinking of the functions as of points in Euclidean space, it is clear that the continuous function $d(f, C)$ attains its minimum on the closed and bounded set $\mathcal{S}$. The uniqueness of $\hat{C}$ follows from the convexity of $\mathcal{S}$ using, e.g., Robertson, Wright, and Dykstra (1988, Theorem 1.3.1). $\square$

2.3 Linear model

With the given option data, constraints 1–3 of Subsection 2.2, can be reformulated using linear regression models with constraints.

In the following, we fix the time t and the expiry date T and we omit these symbols from the notation. In Subsection 2.2 we have noted that the option prices are repeatedly observed for a small number p of distinct strike prices. Defining the expected values of the option prices for a given strike price, $\mu_j = \mu(k_j) = E\{C(k_j)\}$, we can write

$$\begin{aligned}
\mu_p &= \beta_0, \\
\mu_{p-1} &= \beta_0 + \beta_1, \\
\mu_{p-2} &= \beta_0 + 2\beta_1 + \beta_2, \\
\mu_{p-3} &= \beta_0 + 3\beta_1 + 2\beta_2 + \beta_3, \\
&\vdots \\
\mu_1 &= \beta_0 + (p-1)\beta_1 + (p-2)\beta_2 + \cdots + \beta_{p-1}.
\end{aligned}$$

Thus, we fit our data using coefficients β_j, $j = 1, \ldots, p$. The conditional means μ_i, $i = 1, \ldots, p$ are replaced by the same number of parameters β_j, $j = 0, \ldots, p-1$ which allow to impose the shape constraints in a more natural way.

Insert Figure 1

The interpretation of the coefficients β_j can be seen in Figure 1, which shows a simple situation with only four distinct strike prices ($p = 4$). β_0 is the mean option price at point 4. Constraint 1', Subsection 2.1, implies that it has to be positive. β_1 is the difference between the mean option prices at point 4 and point 3; Constraint 2' implies that it has to be positive. The next coefficient, β_2, approximates the change in first derivative in point 3 and it can be interpreted as an approximation of the second derivative in point 3. Constraint 3' implies that β_2 has to be positive. Similarly, β_3 is an estimate of the (positive) second derivative in point 2. Constraint 4' can be rewritten as $\beta_2 + \beta_3 \leq 1$.

In practice, we start with the construction of a design matrix which allows us to write the above model in the following linear form. For simplicity of presentation, we again set $p = 4$:

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}. \tag{5}$$

Ignoring the constraints on the coefficients would lead to a simple linear regression problem. Unfortunately, this approach does not have to lead, and usually does not, to interpretable and stable results.

Model (5) in the above form can be reasonably interpreted only if the observed strike prices are equidistant and if the distances between the neighboring observed strike prices are equal to one. If we want to keep the interpretation of the parameters β_j as the derivatives of the estimated function, we should use the design matrix

$$\Delta = \begin{pmatrix} 1 & \Delta_p^1 & \Delta_{p-1}^1 & \Delta_{p-2}^1 & \cdots & \Delta_3^1 & \Delta_2^1 \\ 1 & \Delta_p^2 & \Delta_{p-1}^2 & \Delta_{p-2}^2 & \cdots & \Delta_3^2 & 0 \\ \vdots & & & & & & \vdots \\ 1 & \Delta_p^{p-2} & \Delta_{p-1}^{p-2} & 0 & \cdots & 0 & 0 \\ 1 & \Delta_p^{p-1} & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \tag{6}$$

where $\Delta_j^i = \max(k_j - k_i, 0)$ denotes the positive part of the distance between k_i and k_j, the i-th and the j-th ($1 \le i \le j \le p$) sorted distinct observed values of the strike price.

The vector of conditional means μ can be written in terms of the parameters β as follows

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \mu = \Delta\beta = \Delta \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{pmatrix}. \tag{7}$$

The constraints on the conditional means μ_j can now be expressed as conditions on the parameters of the model (7). Namely, it suffices to request that $\beta_i > 0$, $i = 0, \dots, p-1$ and that $\sum_{j=2}^{p-1} \beta_j \le 1$.

The model for the option prices can now be written as

$$C(\mathcal{K}) = \mathcal{X}_\Delta \beta + \varepsilon, \tag{8}$$

where $\mathcal{X}_\Delta$ is the design matrix obtained by repeating each row of matrix Δ n_i-times, $i = 1, \ldots, p$.

3 Implementing the constraints

In order to impose Constraints 1–3 on parameters β_i, $i = 0, \ldots, p-1$, we propose the following reparameterization of the model in terms of parameters $\theta = (\theta_0, \ldots, \theta_{p-1})^\top$:

$$\begin{aligned} \beta_0(\theta) &= \exp(\theta_0), \\ \beta_1(\theta) &= \exp(\theta_1), \\ &\vdots \\ \beta_{p-1}(\theta) &= \exp(\theta_{p-1}), \end{aligned}$$

under the constraint that $\sum_{j=2}^{p-1} \exp(\theta_j) < 1$. Clearly, the parameters $\beta_i(\theta)$ satisfy the constraints

$$\begin{aligned} \beta_i(\theta) &> 0, \quad i = 0, \ldots, p-1, \\ \sum_{j=2}^{p-1} \beta_j(\theta) &< 1. \end{aligned}$$

This means that the parameters $\beta_2(\theta), \ldots, \beta_{p-1}(\theta)$ can be considered as point estimates of the state price density (the estimates have to be positive and integrate to less than one). Furthermore, in view of Lemma 1, it is worthwhile to note that the parameters satisfy also

$$-\sum_{j=1}^{k} \beta_j \in (-1, 0), \quad \text{for } k = 1, \ldots, p-1.$$

The model (8) rewritten in terms of parameters θ_i, $i = 0, \ldots, p$ is a nonlinear regression model which can be estimated using standard nonlinear least squares or maximum likelihood methods (Seber and Wild, 2003). The main advantage of these methods is that the asymptotic distribution is well known and that the asymptotic variance of the estimator can be approximated using numerical methods implemented in many statistical packages.

3.1 Reparameterization

The following reparameterization of the model in terms of parameters $\xi = (\xi_0, \ldots, \xi_p)^\top$ simplifies the calculation of the estimates because it guarantees that all constraints are automatically satisfied:

$$
\begin{aligned}
\beta_0(\xi) &= \exp(\xi_0), \\
\beta_1(\xi) &= \frac{\exp(\xi_1)}{\sum_{j=1}^{p} \exp(\xi_j)}, \\
&\vdots \\
\beta_{p-1}(\xi) &= \frac{\exp(\xi_{p-1})}{\sum_{j=1}^{p} \exp(\xi_j)}.
\end{aligned}
$$

This property simplifies the numerical minimization algorithm needed for the calculation of the estimates.

The equality

$$\frac{1}{\sum_{j=1}^{p-1} \beta_j(\xi)} = 1 + \frac{\exp(\xi_p)}{\sum_{j=1}^{p-1} \exp(\xi_j)}$$

shows the meaning of the additional parameter ξ_p. Setting this parameter to $-\infty$ would be the same as requiring that $\sum_{j=1}^{p-1} \beta_j(\xi) = 1$. Large values of the parameter ξ_p indicate that the estimated coefficients sum to less than one or, in other words, the observed strike prices do not cover the support of the estimated SPD. Notice that, by setting $\xi_p = -\infty$, we could easily modify our procedure and impose the equality constraint $\sum_{j=1}^{p-1} \beta_j(\xi) = 1$.

3.2 Inverse transformation of model parameters

For the numerical algorithm, it is useful to know how to calculate ξs from given βs. This is needed, for example, to obtain reasonable starting points for the iterative procedure maximizing the likelihood.

LEMMA 3 *Given $\beta = (\beta_1, \ldots, \beta_p)^\top$, where $\beta_p = 1 - \sum_{i=1}^{p-1} \beta_i$, the parameters $\xi = (\xi_1, \ldots, \xi_p)^\top$ satisfy the system of equations*

$$\left(\beta \mathbf{1}_p^\top - \mathbf{I}_p\right) \exp \xi^\top = \mathcal{A} \exp \xi^\top = 0, \tag{9}$$

where $\mathbf{I}_p$ is the $(p \times p)$ identity matrix. Furthermore,

$$\operatorname{rank} \mathcal{A} = p - 1. \tag{10}$$

The system of equations (9) has infinitely many solutions which can be expressed as

$$\exp(\xi) = \left(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p\right) z, \tag{11}$$

where $\mathcal{A}^{-}$ denotes a generalized inverse of $\mathcal{A}$ and where z is an arbitrary vector in $\mathbb{R}^p$ such that the right hand side of (11) is positive.

Proof:
Parts (9) and (10) follow from the definition of $\beta(\xi)$ and from simple algebra (notice that the sum of rows of $\mathcal{A}$ is equal to zero). Part (11) follows, e.g., from Anděl (1985, Theorem IV.18). □

It remains to choose the vector z in (11) so that the solution of the system of equations (9) is positive.

PROPOSITION 1 *The rank of the matrix $\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p$ is 1. Hence, any solution of the system of equations (9) is a multiple of the first column of the matrix $\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p$. The vector z in (11) can be chosen, e.g., as $z = \pm\mathbf{1}_p$, where the sign is chosen so that the resulting solution is positive.*

Proof:
The definition of a generalized inverse is

$$\mathcal{A}\mathcal{A}^{-}\mathcal{A} - \mathcal{A} = \mathcal{A}(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p) = 0. \tag{12}$$

Lemma 3 says that $\operatorname{rank}\mathcal{A} = p - 1$ and, hence, equation (12) implies that $\operatorname{rank}(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p) \leq 1$. Noticing that $\mathcal{A}^{-}\mathcal{A} \neq \mathbf{I}_p$ means that $\operatorname{rank}(\mathcal{A}^{-}\mathcal{A} - \mathbf{I}_p) > 0$ and concludes the proof. □

3.3 The algorithm

The proposed algorithm consists of the following steps:

1: obtain a reasonable initial estimate $\hat{\beta}$, e.g., by running the Pool-Adjacent-Violators algorithm (Robertson, Wright, and Dykstra, 1988, Chapter 1) on the unconstrained least squares estimates of the first derivative of the curve,
2: transform the initial estimates $\hat{\beta}$ into the estimates $\hat{\xi}$ using the method described in Subsection 3.2,
3: estimate the parameters of the model (8) by minimizing the sum of squares $\{C(\mathcal{K}) - \mathcal{X}_\Delta\beta(\xi)\}^\top\{C(\mathcal{K}) - \mathcal{X}_\Delta\beta(\xi)\}$ in terms of parameters ξ (see Subsection 3.1) using numerical methods.

An application of this simple algorithm on real data is given in Subsection 5.1.

3.4 Asymptotic confidence intervals

We construct confidence intervals based on the parameterization $\beta(\theta)$ introduced at the beginning of this section. The confidence limits for parameters θ_i are exponentiated in order to obtain valid pointwise confidence bounds for the true SPD. The main advantage of this approach is that such confidence bounds are always positive.

An alternative approach would be to construct confidence intervals based on the parameterizations in terms of β_i (Section 2.3) or ξ_i (Section 3.1). However, the limits of confidence intervals for β_i may be negative and confidence intervals for the SPD based on parameters ξ_i would have very complicated shapes in high-dimensional space and could not be easily calculated and interpreted.

Another approach to the construction of the asymptotic confidence intervals can be based on the maximum likelihood theory. Assuming normality, the log-likelihood for the model (8) can be written as:

$$l(\mathcal{C}, \mathcal{X}_\Delta, \theta, \sigma) = -n \log \sigma - \frac{1}{2\sigma^2}\{\mathcal{C} - \mathcal{X}_\Delta \beta(\theta)\}^\top \{\mathcal{C} - \mathcal{X}_\Delta \beta(\theta)\}, \tag{13}$$

where $\mathcal{X}_\Delta$ is the design matrix given in (8). This normality assumption are justified later by a residual analysis. The maximum likelihood estimator is defined as:

$$\hat{\theta} = \arg\max_\theta l(\mathcal{C}, \mathcal{X}_\Delta, \theta, \sigma) \tag{14}$$

and it has asymptotically a p-dimensional normal distribution with mean θ and the variance given by the inverse of the Fisher information matrix:

$$\mathcal{F}_n^{-1} = \left\{ -E\left(\frac{\partial^2}{\partial\theta\partial\theta^\top} l(\mathcal{C}, \mathcal{X}_\Delta, \theta, \sigma) \right) \right\}^{-1}. \tag{15}$$

More precisely, $n^{1/2}(\hat{\theta} - \theta) \xrightarrow{\mathcal{L}} N_p(0, \mathcal{F}_n^{-1})$. In this framework, the Fisher information matrix can be estimated by numerically differentiated Hessian matrix of the log-likelihood. For details we refer, e.g., to Serfling (1980, Chapter 4). The confidence intervals calculated for parameters θ may be transformed (exponentiated) to a confidence intervals for the SPD (β). We have not pursued the maximum likelihood approach since it was numerically less stable in this situation.

Note that, under the assumptions of normality, the maximum likelihood estimate is equal to the nonlinear least squares estimate (Seber and Wild, 2003, Section 2.2) and the asymptotic variance of $\hat{\theta} = \exp(\beta)$ may be approximated by $\operatorname{Var}\widehat{\theta} = \{\operatorname{diag}(\exp\widehat{\theta})\mathcal{X}_\Delta^\top \mathcal{X}_\Delta \operatorname{diag}(\exp\widehat{\theta})\}^{-1}\widehat{\sigma}^2$. Hence, asymptotic confidence intervals for θ_i may be calculated as $(\widehat{\theta}_i \pm u_{1-\alpha/2}\widehat{s}_{ii})$, where $u_{1-\alpha/2}$ is $1-\alpha/2$ quantile of standard Normal distribution and $\widehat{s}_{ii}$ denotes the i-th diagonal element of $\operatorname{Var}\hat{\theta}$. By exponentiating both limits of this confidence interval, we immediately obtain $1-\alpha$ confidence interval for $\beta_i = \exp\theta_i$.

The construction of the estimator guarantees that the matrix $\mathcal{X}_\Delta$ has full rank—this implies that $\mathcal{X}_\Delta^\top \mathcal{X}_\Delta$ is invertible and the asymptotic variance matrix $\operatorname{Var}\widehat{\theta}$ always exists. If the number of observations is equal to the number of distinct strike prices (if there is only one option price for each strike price), it may happen that $\widehat{\sigma}^2 = 0$ and the confidence intervals degenerate to a single point.

3.5 Put-Call parity

The prices of put options can be easily included in our estimation technique by applying the Put-Call parity of the option prices. Assuming that there are no dividends or costs connected with the ownership of the stock, each put option with price $P_t(K,T)$ corresponds to a call option with price

$$C_t(K,T) = P_t(K,T) + S_t - Ke^{-r(T-t)}.$$

In this way, the prices of the put options can be converted into the prices of call options and used in our model (Stoll, 1969). Statistically speaking, these additional observations will increase the precision of the SPD and will lead to more stable results.

In Germany, the Put-Call parity might be biased by an effect of DAX index calculation which is based on the assumption that the dividends are reinvested after deduction of corporate income tax. As the income tax of some investors might be different, the value of DAX has to be corrected before using Put-Call parity in subsequent analysis. For the exact description of this correction we refer to Hafner and Wallmeier (2000) who were analyzing the same data set.

Insert Figure 2

The construction of our estimates allows to include the put option prices in more direct way by fitting the two curves separately using two sets of parameters. The situation is displayed in Figure 2. Our assumption that the same SPD drives both the put and call option prices is naturally translated in terms of the coefficients α_i and β_i

$$\alpha_i = \beta_{p-i+1}, \quad \text{for } i = 2, \dots, p-1$$
$$\alpha_1 = 1 - \sum_{i=1}^{p-1} \beta_i.$$

The problem of estimating regression functions under such linear equality constraints is solved, e.g., in Rao (1973). In Section 4.3, we will investigate also the covariance of the observed call and put option prices and the suggested model will be presented in detail.

3.6 *Expected value constraints*

In Subsection 2.3, we have explained that the parameters $\beta_2, \dots, \beta_{p-1}$ can be interpreted as estimates of the state price density in points $k_2, \dots, k_{p-1}$. From the construction of the estimator, see also Figure 1, it follows that parameter β_1 can be interpreted as the mass of the SPD lying to the right of k_{p-1}. Assuming that the observed strike prices cover entirely the support of the SPD, the mass β_1 could be attributed to the point k_p. Notice that the reparameterization introduced in Subsection 3 guarantees that $\sum_{i=1}^{p-1} \beta_i(\xi) < 1$ and it immediately follows that interpreting β_1 as the estimate of the SPD in point k_p does not violate any constraints described in Subsection 2.2.

Referring to the previous Subsection 3.5 it is clear that the parameter $\beta_p \equiv \alpha_1 = 1 - \sum_{i=1}^{p-1} \beta_i$ can be interpreted as the estimator of the SPD in k_1. The parameterization of the problem now guarantees that $\sum_{i=1}^{p} \beta_i = 1$.

The expected value of the underlying stock under the risk-neutral measure can now be estimated as $\widehat{E^{\text{SPD}}} = \sum_{i=1}^{p} k_i \beta_{p-i+1}$. From economic theory it follows that $\widehat{E^{\text{SPD}}}$ has to be equal to the forward price of the stock. This constraint can be easily implemented by using the fact that β_1 and β_p estimate the mass of the SPD respectively to the right of k_{p-1} and to the left of k_2.

If $\widehat{E^{\text{SPD}}}$ is smaller than the forward price $\exp\{r(T-t)\}S_t$ of the stock, it suffices to move the mass β_1 further to the right. If $\widehat{E^{\text{SPD}}}$ is too large, we move the mass β_p to the left. More precisely, setting

$$\widetilde{k}_1 = k_1 - \boldsymbol{I}(\widehat{E^{\text{SPD}}} > \exp\{r(T-t)\}S_t)(\widehat{E^{\text{SPD}}} - \exp\{r(T-t)\}S_t)/\beta_p,$$
$$\widetilde{k}_p = k_p + \boldsymbol{I}(\widehat{E^{\text{SPD}}} < \exp\{r(T-t)\}S_t)(\exp\{r(T-t)\}S_t - \widehat{E^{\text{SPD}}})/\beta_1,$$

we get

$$\exp\{r(T-t)\}S_t = \widetilde{k}_1\beta_p + \sum_{i=2}^{p-1} k_i\beta_{p-i+1} + \widetilde{k}_p\beta_1.$$

This choice of $\widetilde{k}_1$ and $\widetilde{k}_p$ guarantees that the expected value corresponding to the estimator $\beta_1, \ldots, \beta_p$ is equal to the forward price S_t of the stock, see the beginning of Section 6 for an application of this technique.

In the following Sections 5 and 4, we will concentrate on the properties of $\beta_2, \ldots, \beta_{p-1}$ and further improvements in the estimation procedure.

4 Covariance structure

In this section, we use a model for the SPD development throughout the day to derive the covariance structure of the observed option prices depending on the strike prices and time of the trade. Considering the covariance structure in the estimation procedure solves the problems with heteroscedasticity and correlation of residuals that will be demonstrated in Subsection 5.3.

In this model, most recent option prices have the smallest variance and thus the largest weight in the estimation procedure. Similarly, the covariance of two option prices with the same strike price at approximately the same time is larger than covariances of prices of some more dissimilar options.

We start by rewriting the model with iid error terms so that it can be more easily generalized. In Subsection 4.1, we present a model that accounts for heteroscedasticity and which is further developed in Subsections 4.2 and 4.3 where an approximation of the covariance is calculated for any two options prices using only their strike prices and time of the trade. In Subsection 4.4, we suggest to decompose the error term into two parts and we show how to estimate these additional parameters by maximum likelihood method. The analysis of the resulting standardized residuals in Subsection 5.4 suggests that this covariance structure is applicable to our dataset.

Until now, we assumed that the i-th option price (on a fixed day t) satisfies

$$C_i(k_j) = \Delta_j \widetilde{\beta} + \varepsilon_i \tag{16}$$

or

$$\begin{aligned} C_i(k_j) &= \Delta_j \widetilde{\beta}_i + \varepsilon_i, \\ \widetilde{\beta}_i &= \widetilde{\beta}_{i-1}, \end{aligned} \tag{17}$$

where ε_i are iid random errors with zero mean and constant variance σ^2, $\widetilde{\beta} = \widetilde{\beta}_1 = \cdots = \widetilde{\beta}_i$ denotes the column vector of the unknown parameters, and

Δ_j denotes the j-th row of the matrix Δ defined in (6), i.e.,

$$\Delta_j = (1, \Delta^j_p, \Delta^j_{p-1}, \dots, \Delta^j_{j+1}, \underbrace{0, \dots, 0}_{(j-1)}).$$

The residual analysis in Section 5.3 clearly demonstrates that the random errors ε_i are not independent and homoscedastic and we have to consider some generalizations that lead to a better fit of the data set.

4.1 Heteroscedasticity

Assume that the i-th observation, corresponding to the j-th smallest exercise price k_j, can be written as

$$C_i(k_j) = \Delta_j \widetilde{\beta}_i, \tag{18}$$

$$\widetilde{\beta}_i = \widetilde{\beta} + \varepsilon_i, \tag{19}$$

i.e., there are iid random vectors ε_i having iid components with zero mean and variances σ^2 in the state price density $\widetilde{\beta}_i$. Clearly, the variance matrix of the vector of the observed option prices C is then

$$\operatorname{Var} C = \sigma^2 \operatorname{diag}(\mathcal{X}_\Delta \mathcal{X}_\Delta^\top), \tag{20}$$

where $\mathcal{X}_\Delta$ is the design matrix in which each row of the matrix Δ is repeated n_j times, $j = 1, \dots, p$.

REMARK 2 *Assuming that the observed option prices have the covariance structure (20), the least squares estimates do not change and*

$$\operatorname{Var} \hat{\beta} = \sigma^2 \{\mathcal{X}_\Delta^\top \operatorname{diag}(\mathcal{X}_\Delta \mathcal{X}_\Delta^\top)^{-1} \mathcal{X}_\Delta\}.$$

Another possible model for the heteroscedasticity would assume that the changes are multiplicative rather than additive.

$$C_i(k_j) = \Delta_j \widetilde{\beta}_i$$

$$\log \widetilde{\beta}_i = \log \widetilde{\beta} + \varepsilon_i$$

This model leads to a variance of $C_i(k_j)$ that depends on the value of the SPD:

$$\operatorname{Var} C_i(k_j) = \sigma^2 \{\beta_0^2 + (\Delta^j_p)^2 \beta_1^2 + (\Delta^j_{p-1})^2 \beta_2^2 + (\Delta^j_{p-2})^2 \beta_3^2 + \dots + (\Delta^j_{j+1})^2 \beta_j^2\}.$$

It is straightforward that Remark 2 applies also in this situation.

4.2 Covariance

Let us now assume that there are random changes in the state price density coefficients $\widetilde{\beta}_i$ over time so that we have

$$\begin{aligned} C_i(k_j) &= \Delta_j \widetilde{\beta}_i, \\ \widetilde{\beta}_i &= \widetilde{\beta}_{i-1} + \varepsilon_i, \end{aligned} \tag{21}$$

where, for fixed i, $\widetilde{\beta}_i$ is the parameter vector and ε_k, $k = i, i-1, \ldots$ are iid random vectors having iid components with zero mean and variances σ^2. For nonequidistant time points, let δ_i denote the time between the i-th and $(i-1)$-st observation. The model is

$$\begin{aligned} C_i(k_j) &= \Delta_j \widetilde{\beta}_i, \\ \widetilde{\beta}_i &= \widetilde{\beta}_{i-1} + \delta_i^{1/2} \varepsilon_i \end{aligned} \tag{22}$$

and it leads to the covariance matrix with elements

$$\begin{aligned} \operatorname{Cov}\{C_{i-u}(k_j), C_{i-v}(k_i)\} &= \operatorname{Cov}(\Delta_j \widetilde{\beta}_{i-u}, \Delta_i \widetilde{\beta}_{i-v}) \\ &= \sigma^2 \Delta_j \Delta_i^\top \sum_{l=1}^{\min(u,v)} \delta_{i+1-l}. \end{aligned} \tag{23}$$

When we observe the i-th observation, we are usually interested in the estimation of the current value of the vector of parameters $\widetilde{\beta}_i$.

4.3 Including put options

Similarly, we obtain the covariance for the price of the put options, $P_i(k_j)$. Using the relations between the α and β parameters, $\alpha_k = \beta_{p-k+1}$, for $k = 2, \ldots, p-1$ and after some simplifications, we can write the model for the price of the put options, $P_i(k_j)$, as

$$\begin{aligned} P_i(k_j) &= \Delta_j \widetilde{\alpha}_i, \\ \widetilde{\alpha}_i &= \widetilde{\alpha}_{i-1} + \delta_i^{1/2} \varepsilon_i, \end{aligned} \tag{24}$$

where $\widetilde{\alpha} = (\alpha_0, \alpha_1, \beta_{p-1}, \beta_{p-2}, \ldots, \beta_2)^\top$ and Δ_j^P denotes the corresponding row of the design matrix, i.e.,

$$\Delta_j^P = (1, \Delta_j^1, \Delta_j^2, \ldots, \Delta_j^{j-1}, \underbrace{0, \ldots, 0}_{(p-j)}).$$

In this way, we obtain a joint estimation strategy for both the call and put option prices:

$$
\begin{aligned}
&C_i(k_j) = \Delta_j \widetilde{\beta}_i,\\
&P_i(k_j) = \Delta_j^P \widetilde{\alpha}_i,\\
&\begin{pmatrix} \widetilde{\beta}_i \\ \widetilde{\alpha}_i \end{pmatrix} = \begin{pmatrix} \widetilde{\beta}_{i-1} \\ \widetilde{\alpha}_{i-1} \end{pmatrix} + \delta_i^{1/2} \varepsilon_i,
\end{aligned} \tag{25}
$$

which directly leads to covariance

$$
\begin{aligned}
\mathrm{Cov}\{P_{i-u}(k_j), P_{i-v}(k_i)\} &= \mathrm{Cov}(\Delta_j^P \widetilde{\alpha}_{i-u}, \Delta_i^P \widetilde{\alpha}_{i-v})\\
&= \sigma^2 \Delta_j^P (\Delta_i^P)^\top \sum_{l=1}^{\min(u,v)} \delta_{i+1-l}
\end{aligned} \tag{26}
$$

and

$$
\begin{aligned}
&\mathrm{Cov}\{C_{i-u}(k_j), P_{i-v}(k_i)\}\\
&\qquad = \mathrm{Cov}(\Delta_j \widetilde{\beta}_{i-u}, \Delta_i^P \widetilde{\alpha}_{i-v})\\
&\qquad = \sigma^2 \sum_{l=1}^{\min(u,v)} \delta_{i+1-l} \sum_{k=2}^{p-1} \Delta_{p+1-k}^{j} \Delta_i^{p+1-k}.
\end{aligned} \tag{27}
$$

Together with (23), equations (26) and (27) allow to calculate the covariance matrix of all observed option prices using only their strike prices and the times between the transactions.

4.4 Error term for option prices

Using the model (25) would mean that all changes observed in the option prices are due only to changes in the SPD. It seems natural to add another error term, η_i, as a description of the error in the option price:

$$
\begin{aligned}
&C_i(k_j) = \Delta_j \widetilde{\beta}_i + \eta_i,\\
&P_i(k_j) = \Delta_j^P \widetilde{\alpha}_i + \eta_i,\\
&\begin{pmatrix} \widetilde{\beta}_i \\ \widetilde{\alpha}_i \end{pmatrix} = \begin{pmatrix} \widetilde{\beta}_{i-1} \\ \widetilde{\alpha}_{i-1} \end{pmatrix} + \delta_i^{1/2} \varepsilon_i,
\end{aligned} \tag{28}
$$

where $\eta_i \sim N(0, \nu^2)$ are iid random variables independent of the random vectors ε_i. Here, normality assumptions are added both for η_i and ε_i so that the variance components parameters ν^2 and σ^2 may be estimated by the maximum likelihood method.

Next, in order to simplify the notation, let us fix the index i and let Y denote the vector of observed call and put option prices, $\mathcal{X}_\Delta$ the corresponding design matrix consisting of the corresponding rows Δ_j and Δ_j^P, and $\widetilde{\gamma}$ the combined vector of unknown parameters. Denoting by Σ_i the matrix containing the covariances defined in (23,26,27), we can rewrite the model (25) as

$$Y = \mathcal{X}_\Delta \widetilde{\gamma} + \xi, \tag{29}$$

where $\operatorname{Var} \xi = \operatorname{Var} Y = \sigma^2 \Sigma_i + \nu^2 \mathcal{I}_n = \sigma^2(\Sigma_i + \psi^2 \mathcal{I}_n) = \sigma^2 V$, where $\psi^2 = \nu^2/\sigma^2$. Differentiating the log-likelihood

$$l(\beta, \sigma^2, \psi^2) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\sigma^2 V| - \frac{1}{2\sigma^2}(Y - \mathcal{X}_\Delta\widetilde{\gamma})^\top V^{-1}(Y - \mathcal{X}_\Delta\widetilde{\gamma})$$

we obtain

$$\frac{\partial l(\beta, \sigma^2, \psi^2)}{\partial \psi^2} = -\frac{1}{2}\operatorname{tr}(V^{-1}) + \frac{1}{2\sigma^2}(Y - \mathcal{X}_\Delta\widetilde{\gamma})^\top V^{-2}(Y - \mathcal{X}_\Delta\widetilde{\gamma}). \tag{30}$$

For any fixed value of the parameter ψ^2, it is straightforward to calculate optimal σ^2 and $\widetilde{\gamma}$. Hence, the numerical maximization of the log-likelihood can be based on a search for a root (zero) of the one-dimensional function (30).

Moreover, the variance components parameters σ^2 and $\nu^2 = \psi^2\sigma^2$ have a very natural econometric interpretation: σ^2 describes the speed of change of the SPD and ν^2 the error in observed option prices.

5 Application to DAX data

We analyze a data set containing observed option prices for various strike prices and maturities. Other variables are the interest rate, date, and time. In 1995, one observed every day about 500 trades, in todays more liquid option markets this number has increased approximately 10 times. In our empirical study we will consider the time period from 1995 up to 2003 thus also covering more recent liquid option market.

Figure 3 displays the observed prices of European call options written on the DAX for the 16. January 1995. The left panel shows the ensemble of call option prices for different strikes and maturities as a free structure together with a smooth surface. The typical shape of dependency of the option price on the strike price can be observed on the right panel containing the option prices only for the shortest time to expiry, $\tau = T - t = 4$ days.

Insert Figure 3

In order to illustrate the method, we apply it to DAX option prices on two consecutive days. These days (16th and 17th January 1995) were selected since they provide nice insight into the behavior of the presented methods.

5.1 Estimator with iid random errors

We start by a comparison of the unconstrained and constrained estimator described respectively in Subsections 2.3 and 3.1.

Insert Figure 4

For the European call option prices displayed in the right hand plot in Figure 3, we obtain the estimates plotted in Figure 4. The top plot displays the original data, the second plot shows the estimate of the first derivative, and the third plot shows the estimate of the second derivative, i.e., the state price density. Actually, all plots contain two curves, both obtained using model (8). The thick line is calculated using the parameters β_i without constraints whereas the thin line uses the reparameterization $\beta_i(\xi)$ given in Subsection 3.1. In Figure 4, these two estimates coincide since the model maximizing the likelihood without constraints, by chance, fulfills the constraints ($\exists \xi : \beta_i = \beta_i(\xi)$, $i = 0, \ldots, p-1$) and hence it is clear that the same parameters maximize also the constrained likelihood.

Insert Figure 5

The situation, in which the call pricing functions fitted with and without constraints differ, is displayed in Figure 5. Notice that the difference between the two regression curves is small whereas the difference between the estimates of the state price density (i.e., the second derivative of the curve) is surprisingly large. The unconstrained estimate shows very unstable behavior on the left hand side of the plot. The constrained version behaves more reasonably. Very small differences between the fitted call pricing functions in the top plot in Figure 5 leads to huge differences in the estimates of second derivative.

We therefore conclude that small error in the estimate of the call pricing function may lead to large scale error in the estimates of the first and second derivatives. The scale of this type of error seems to be limited by imposing the shape constraints given in Subsection 2.2.

5.2 Confidence intervals

Insert Figure 6

Insert Figure 7

In Figures 6 and 7 we plot both estimates together with 95% confidence intervals. Notice that, in the unconstrained model, the estimates of the values of the SPD are just the parameters of the linear regression model. Hence, the confidence intervals for the parameters are, at the same time, also confidence intervals for the SPD. These confidence intervals for 16th and 17th January are displayed in the upper plots in Figures 6 and 7. The drawbacks of this method are clearly visible. In Figure 6, the lower bounds of the confidence intervals only asymptotically satisfy the condition of positivity. In Figure 7, we observe large variability on the left-hand side of the plot (the region with low number of observations). Again, some of the lower bounds are not positive. Clearly, the confidence intervals based on the unconstrained model make sense only if the constraints are, by chance, satisfied. Even if this is the case, there is no guarantee that the lower bounds will be positive. The lower panels in Figures 6 and 7 display the nonnegative asymptotic confidence intervals calculated according to Subsection 3.4.

In Figure 6, both type of confidence intervals provide very similar results. The only difference is at the minimum and maximum value of the independent variable (strike price) where the unconstrained method provides negative lower bounds and the conditional method leads to very large upper bounds of the confidence intervals.

In Figure 7, we plot the confidence intervals for January 17th. In the central region of the graphics, both types of confidence intervals are quite similar. On the left and right hand side, both methods tend to provide confidence intervals that seem to be overly wide. For the constrained method, we observe that the length of the confidence intervals explodes when the estimated value of the SPD is very close to zero and, at the same time, the number of observation in that region (see the description of the horizontal axis) is small.

5.3 Residual analysis

Insert Figure 8

The residuals on 17th January 1995 are plotted in Figure 8. The time of trade (in hours) is denoted by the plotting symbol. The circle, square and the star denote denote the trades carried out in the morning, midday and in the afternoon, respectively. The size of the symbols corresponds to the number of residuals lying in the respective areas.

The majority of the residuals correspond to the strike prices of 2075DEM and 2100DEM. The variance of the residuals is very low on the right hand side of the plot and it rapidly increases when moving towards smaller strike prices. On the left hand side of the plot, for strike prices smaller than 2000, we have only very few observations and cannot judge the residual variability reliably.

Apart of the obvious heteroscedasticity we observe also a very strong systematic movement in the SPD throughout the day: the circles, corresponding to the first third of the day, are positive and all stars, denoting the afternoon residuals, are negative. Similar patterns can be observed every day—residuals corresponding to the same time are having the same sign.

We conclude that the assumption of iid random errors is obviously not fulfilled as the option prices tend to follow the changes of the market during the day.

5.4 Application of the covariance structure

Insert Figure 9

In Figure 9, we present the estimator combining both put and call option prices and using the covariance structure proposed in Subsection 4.4. In comparison with the results plotted in Figure 7, we observe shorter length of the confidence intervals.

The estimates of the variance components parameters are $\hat{\psi}^2 = 17.77$, $\hat{\sigma}^2 = 0.0041$, and $\hat{\nu}^2 = 0.0722$. For interpretation, it is more natural to consider $\hat{\nu} = 0.2687$ suggesting that 95% of the option prices were on 17th January 1995 not further than 0.5DEM from the correct option price implied by the current (unobserved) SPD.

Insert Figure 10

The standardized residuals in the top panel of Figure 10 were plotted using the same technique as the residuals in Figure 8. Whereas the residuals for the iid model showed strong correlations and heteroscedasticity, the structure of the standardized residuals looks much better. It is natural that the residuals are larger in the central part since more than 90% of observations have strike price between 2050 and 2100. The largest residuals were omitted in the residual plot so that the structure in the central part is more visible but the lower panel of Figure 10 displays the histogram of all residuals. The distribution of the residuals seems to be symmetric and its shape is not too far from Normal distribution. However, the kurtosis of this distribution is too large and formal tests reject normality.

Insert Figure 11

In Figure 11 we plot prediction intervals for the SPD obtained only by recalculating the covariance structure (28) with respect to some future time. More precisely, the prediction intervals are obtained from option prices observed until i. Then, using the notation of Subsection 4.4, we have for the future $\widetilde{\beta}_{i+1}$ and $\widetilde{\alpha}_{i+1}$:

$$\begin{aligned} C_i(k_j) &= \Delta_j \widetilde{\beta}_i + \eta_i, \\ P_i(k_j) &= \Delta_j^P \widetilde{\alpha}_i + \eta_i, \\ \begin{pmatrix} \widetilde{\beta}_{i+1} \\ \widetilde{\alpha}_{i+1} \end{pmatrix} &= \begin{pmatrix} \widetilde{\beta}_i \\ \widetilde{\alpha}_i \end{pmatrix} + \delta_{i+1}^{1/2} \varepsilon_{i+1}. \end{aligned} \tag{31}$$

It is now easy to see that the only modification that has to be done for estimating $\widetilde{\beta}_{i+1}$ is to add the length of the forecasting horizon δ_{i+1} to the sum in (23), (26), and (27) and to recalculate the confidence regions using this variance matrix with the same estimates of the variance parameters σ^2 and ν^2. In Figure 11, the 95% confidence intervals for the true SPD are denoted by black dashed line. The grey dashed lines denote the prediction intervals calculated for each 30 minutes for next 5 hours. In this way, we can obtain a simple approximation for future short-term fluctuations of the SPD. In long run, the prediction intervals become too wide to be informative.

6 Dynamics of SPD

In order to study the dynamics of SPDs, we calculated basic moment characteristics of the estimated SPDs. Note that the estimator does not allow to estimate the SPD in the tails of the distribution. We can only estimate the probability mass lying to the left $(1 - \sum_{i=1}^{p-1} \beta_i)$ and to the right (β_1) of the available strike price range. Hence, the moments calculated in this section are only approximations which cannot be calculated more precisely without additional assumptions, e.g., on the tail behavior or parametric shape of the SPD.

Insert Figure 12

Insert Figure 13

The estimated mean and variance in the first Quarter of 1995 are plotted as lines in Figures 12–13. Note, that the SPDs in this period were always estimated using the options with shortest time to maturity. This means that the time to maturity is decreasing linearly in both plots, but it jumps up whenever the option with the shortest time to maturity expires. These jumps occurred at days 16, 36, and 56.

From no-arbitrage considerations it follows that the mean of the SPD should correspond to the value of DAX,

$$\widehat{E^{\mathrm{SPD}}} = \int S_T f(S_T) dS_T = \exp\{r(T-t)\} S_t.$$

see also the discussion in Subsection 3.6. In Figure 12, the observed values of DAX multiplied by the factor $\exp\{r(T-t)\}$ are plotted as circles for the first 65 trading days in 1995 and we observe that the estimated means of the SPD estimates, displayed as the line, follow very closely the theoretical value. A small difference is mainly due to the fact that in 1995, the observed strike prices do not cover entirely the support of the SPD. For example, on day 16, the difference between the SPD mean (2018.7) and the DAX multiplied by the discount factor (2012.1) is equal to 6.6. The fact that there are not any trades for strike prices smaller than 1925 means that we only know that the probability mass lying to the left from 1950 is equal to 0.25. In the calculation of the estimate of the SPD mean plotted in Figure 12, this probability mass is assigned to the value 1925 as this is the leftmost observed strike price. Obviously, assigning this probability mass rather to the value $1925 - (6.6/0.25) = 1898.6$ leads more realistic estimate of the SPD and to the equality of the SPD mean and the discounted DAX.

In Figure 13, we see that the variance of SPD decreases linearly as the option moves closer to its maturity. This observation suggests that SPD estimates calculated for neighboring maturities can be linearly interpolated in order to obtain SPD estimate with arbitrary time to maturity. Such an estimate is important for making the SPD estimates comparable and for studying the development of the market expectations.

6.1 Estimate with the fixed time to expiry

The variances displayed in Figure 13 suggest that the variance of the SPD estimates changes approximately linearly in time when moving closer to the date of expiry.

Hence, from the estimates $f_{\tau_1}(.)$ and $f_{\tau_2}(.)$ of centered SPDs corresponding to the times of expiry $\tau_1 < \tau_2$, we construct an estimate $f_\tau(.)$ for any time of expiry $\tau \in (\tau_1, \tau_2)$ as

$$f_\tau(.) = \frac{(\tau_2 - \tau) f_{\tau_1}(.) + (\tau - \tau_1) f_{\tau_2}(.)}{\tau_2 - \tau_1}. \tag{32}$$

In this way, the variance, V_τ, of the centered SPD with time to expiry equal to τ can be expressed as

$$\begin{aligned} V_\tau &= \int x^2 f_\tau(x) dx \\ &= \int x^2 \frac{(\tau_2 - \tau) f_{\tau_1}(x) + (\tau - \tau_1) f_{\tau_2}(x)}{\tau_2 - \tau_1} dx \\ &= \frac{(\tau_2 - \tau) V_{\tau_1} + (\tau - \tau_1) V_{\tau_2}}{\tau_2 - \tau_1}. \end{aligned}$$

We argue that such an estimate is reasonable since we observed in Figure 13 that the SPD variances change linearly in time.

6.2 Verification of the market's expectations

Insert Figure 14

Under the risk neutral (equivalent martingale) measure, the SPD reflects market's expectation of the behavior of the value of the DAX in 45 days. Hence, it is interesting to use our data set to verify how these expectations compare with reality. In the left plot in Figure 14, we plot intervals based on the SPD together with the true future value of the DAX: the black lines display the 2.5% and 97.5% quantiles of the estimated SPD, the future value of DAX is displayed as a grey line. In the right plot, we show in the same way the 45-days ahead predictions based on the historical distribution of the 45-days absolute returns in the last 100 trading days, the 2.5% and 97.5% quantiles of this distribution are plotted as black lines.

Figure 14 suggests that the method works well and that the DAX mostly stays well within the quantiles calculated from the estimated SPDs. The DAX was sometimes rising faster than the market expected from 1995 till mid 1998. After a fast decrease in the second half of 1998, the market increases again till the beginning of year 2000. Since then, the market decreases. However, the changes stay mostly within or very close to the bounds predicted by our SPD estimates. The only exception is a large shock observed in September 2001 caused by the terrorist attack on the World Trade Center.

The upper quantiles, 97.5%, of the historical distribution of 45-days absolute returns mostly agree with the upper quantiles of the SPD. The lower quantiles, 2.5%, of the SPDs seem to be much more variable than the same quantiles of the historical distribution. Both the lower and the upper quantiles of the historical distribution lie mostly above the corresponding quantiles of the estimated SPD, respectively in 69.44% and 81.75%. This observation just confirms the fact that the observed SPD includes effects of risk aversion.

Insert Table 2

In Table 2, we show the fraction of the year that the DAX stays in the prediction corridor. This suggests that the coverage is slightly better for the historical simulation if the DAX is increasing and better for the SPD based prediction if DAX is decreasing (years 2000 and 2002).

6.3 Evaluation of the quality of the forecasts

The quality of the forecasts can be evaluated by comparing the true future observation with its predicted distribution (the SPD). Diebold, Gunther, and Tay (1998) propose to evaluate density forecasts using the probability integral transformed observations $z_{h,t}$, where t denotes the time and h the forecasting horizon. More precisely, we define

$$z_{h,t} = \int_{-\infty}^{X_{t+h}} \widehat{f}_{h,t}(u)du,$$

where $\widehat{f}_{h,t}(.)$ denotes our estimate of the SPD h days ahead at time t and X_{t+h} is the future observation. In other words, $z_{h,t}$ is the probability value of X_{t+h} with respect to $\widehat{f}_{h,t}(.)$. Clearly, the $z_{h,t}$ should be uniformly $U(0,1)$ distributed if the estimated SPD $\widehat{f}_{h,t}(.)$ is equal to the true density of X_{t+h}. In Figure 15, we display the histograms of $z_{h,t}$'s for each year for the estimated SPDs and historical simulation using full and dashed histograms, respectively. Clearly, in the ideal case, the histograms should not be too far from Uniform $U(0,1)$ distribution. In our data, for the prediction horizon $h = 45$ days, we observe that the histograms look quite different than we would expect. Especially in years 1995–1999, the DAX was moving mainly in the upper quantiles of the predicted SPD. The forecasts based on the historical distribution of the 45-days returns behave similarly.

Insert Figure 15

In order to account for the overlapping forecasting periods, we calculate the confidence limits for the empirical distribution function

$$\widehat{F}(u) = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{I}(z_{h,t} \leq u)$$

of $z_{h,t}$'s that take into account the autocorrelation structure.

$$\widehat{\mathrm{Var}}\{\widehat{F}(u)\} = \frac{1}{T}\left\{\widehat{\gamma}_u(0) + 2\sum_{j=1}^{h}\left(1 - \frac{j}{T}\right)\widehat{\gamma}_u(j)\right\}, \tag{33}$$

where $\gamma_u(j)$ is the sample autocovariance of order j:

$$\gamma_u(j) = \frac{1}{T}\sum_{t=j+1}^{T}\left\{\boldsymbol{I}(z_{h,t} \le u) - \widehat{F}(u)\right\}\left\{\boldsymbol{I}(z_{h,t-j} \le u) - \widehat{F}(u)\right\}.$$

Insert Figure 16

The empirical distribution functions $\widehat{F}(.)$ are plotted separately for years 1995–2002 in Figure 16. The distribution function of $U(0,1)$ and the limits following from (33) are displayed as dotted lines. The year 2003 was not included since our dataset contains only two months of the year 2003 which did not leave enough observations to confirm the forecasts.

In years 1996 and 1997, the market was growing much faster than the SPDs were indicating. In year 1996, it never happened that the DAX fell below the 10% quantile of the SPD and there were only few days when this value was below 20%. The situation in years 1998 and 1999 was less extreme even though the fast growth of DAX continued. The distribution given by the SPD estimate $\widehat{f}_{t,h}(.)$ for the horizon $h = 45$ days does not differ significantly from the true distribution of X_{t+h} in years 2000–2001 but in 2002, we again observe significant differences. Thus, the DAX was growing faster than the option market expected in years 1996, 1997, and 1999 and it was falling faster in 2002.

Insert Figure 17

Figure 17 shows the same graphics for forecast based on the historical distribution of the returns. The deviations are more clearly visible but the overall picture is very similar, the only difference arises in 2001 when the predictions did not stay between the limits.

7 Conclusion

We have proposed a simple nonparametric model for arbitrage free estimation of the SPD. Our procedure takes care of the daily changing covariance structure and involves both types of European options. Moreover, the covariance structure allows to calculate prediction intervals capturing future behavior of the SPD. We analyze the moment dynamics of the SPD from 1995–2003. An application to DAX EUREX data for the years 1995–2003 produces a corridor that is compared to the future DAX index value. The proposed technique enables not only to price exotic options but also to measure the risk and volatility ahead of us.

8 Acknowledgments

We thank Volker Krätschmer for useful comments concerning the existence and uniqueness of the constrained regression function and the anonymous referee for many insightful comments leading to substantial improvements in both the presentation and the content of the paper. The research was supported by Deutsche Forschungsgemeinschaft, SFB 649 "Ökonomisches Risiko", by MSM0021620839, GAČR GA201/08/0486, and by MŠMT 1K04018.

References

Aït-Sahalia, Y. and J. Duarte, 2003, Nonparametric option pricing under shape restrictions. Journal of Econometrics 116, 9–47.

Aït-Sahalia, Y. and A.W. Lo, 1998, Nonparametric estimation of state-price densities implicit in financial asset prices. Journal of Finance 53, 499–547.

Aït-Sahalia, Y. and A.W. Lo, 2000, Nonparametric risk management and implied risk aversion. Journal of Econometrics 94, 9–51.

Aït-Sahalia, Y., Y. Wang, and F. Yared, 2000, Do option markets correctly price the probabilities of movement of the underlying asset? Journal of Econometrics 102, 67–110.

Anděl, J., 1985, Mathematical statistics, in Czech. SNTL/Alfa, Prague.

Andersen, L.B.G. and R. Brotherton-Ratcliffe, 1997, The equity option volatility smile: An implicit finite-difference approach. Journal of Computational Finance 1(2), 5–37.

Bondarenko, O., 2003, Estimation of risk-neutral densities using positive convolution approximation. Journal of Econometrics 116, 85–112.

Breeden, D. and R. Litzenberger, 1978, Prices of state-contingent claims implicit in option prices. Journal of Business 51, 621–651.

Buehler, H., 2006, Expensive martingales. Quantitative Finance 6(3), 207–218.

Diebold, F.X., T. Gunther, and A. Tay, 1998, Evaluating density forecasts, with applications to financial risk management. International Economic Review 39, 863–883.

Dupire, B., 1994, Pricing with a smile. RISK 7(1), 18–20.

Fengler, M.R., 2005, Semiparametric modeling of implied volatility. Springer, Heidelberg.

Fengler, M.R., W. Härdle, and E. Mammen, 2007, A dynamic semiparametric factor model for implied volatility string dynamics. Journal of Financial Econometrics 5(2), 189–218.

Hafner, R. and M. Wallmeier, 2000, The Dynamics of DAX Implied Volatilities, University of Augsburg Working Paper. Available at SSRN: http://ssrn.com/abstract=234829 or DOI: 10.2139/ssrn.234829.

Harrison, J. and S. Pliska, 1981, Martingale and stochastic integral in the theory of continuous trading. Stochastic Processes and their Applications 11, 215–260.

Hlávka, Z. and M. Svojík, 2008, Application of extended Kalman filter to SPD estimation, in W. Härdle, N. Hautsch, L. Overbeck, (Eds.), Applied quantitative finance, Springer, Berlin, pp. 233-247.

Huynh, K., P. Kervella, and J. Zheng, 2002, Estimating state-price densities with nonparametric regression, in W. Härdle, T. Kleinow and G. Stahl, (Eds.), Applied quantitative finance. Springer, Heidelberg, pp. 171–196.

Jackwerth, J.C., 1999, Option-implied risk-neutral distributions and implied binomial trees: a literature review. Journal of Derivatives 7, 66–82.

Kahalé, N., 2004, An arbitrage-free interpolation of volatilities. RISK 17(5), 102–106.

Rao, C.R., 1973, Linear statistical inference and its applications. Wiley, New York.

Renault, E., 1997, Econometric models of option pricing errors, in D.M. Kreps and K.F. Wallis, (Eds.), Advances in economics and econometrics: theory and applications, seventh world congress, Vol. III. Cambridge University Press, Cambridge, pp. 223–278.

Robertson, T., F.T. Wright, and R.L. Dykstra, 1988, Order restricted statistical inference. Wiley, Chichester.

Seber, G.A.F. and C.J. Wild, 2003, Nonlinear regression. Wiley, Hoboken, New Jersey.

Serfling, R., 1980, Approximation theorems of mathematical statistics. Wiley, New York.

Stoll, H.R., 1969, The relationship between put and call option prices. Journal of Finance 24, 801–824.

Yatchew, A. and W. Härdle, 2006, Nonparametric state price density estimation using constrained least squares and the bootstrap. Journal of Econometrics 133(2), 579–599.

	Methods			
	Parametric	*Standard smoothing method*	*Nonparametric under constraints*	*This paper*
Shape	fixed	flexible	flexible	flexible
Control	choice of family	smoothness	smoothness	none
SPD support	infinite	restricted	restricted	restricted
Constraints	by design	local	yes	yes

Table 1
Summary of properties of parametric and nonparametric estimators.

Year	1996	1997	1998	1999	2000	2001	2002
SPD	84.40%	66.13%	75.30%	74.60%	97.22%	85.66%	94.84%
Historical	82.00%	79.44%	76.89%	77.38%	93.25%	86.06%	80.56%

Table 2
Fraction of the year that DAX stays in the prediction corridor.

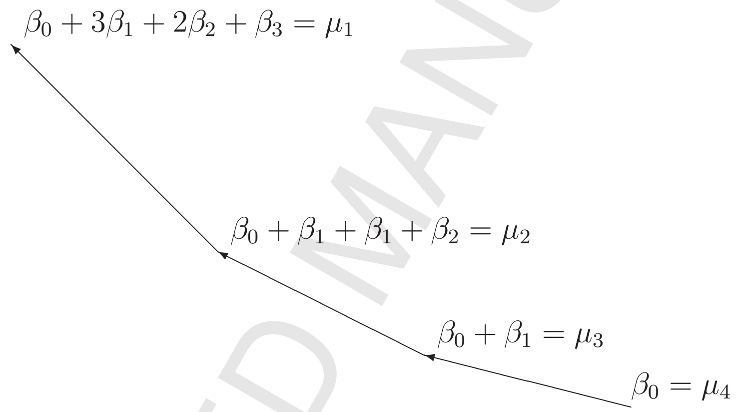


Fig. 1. Illustration of the dummy variables for call options.

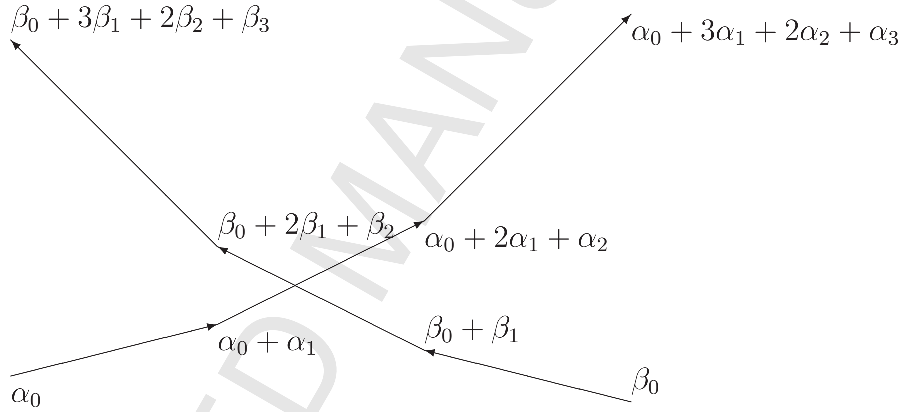


Fig. 2. Illustration of the dummy variables for both call (β) and put (α) options.

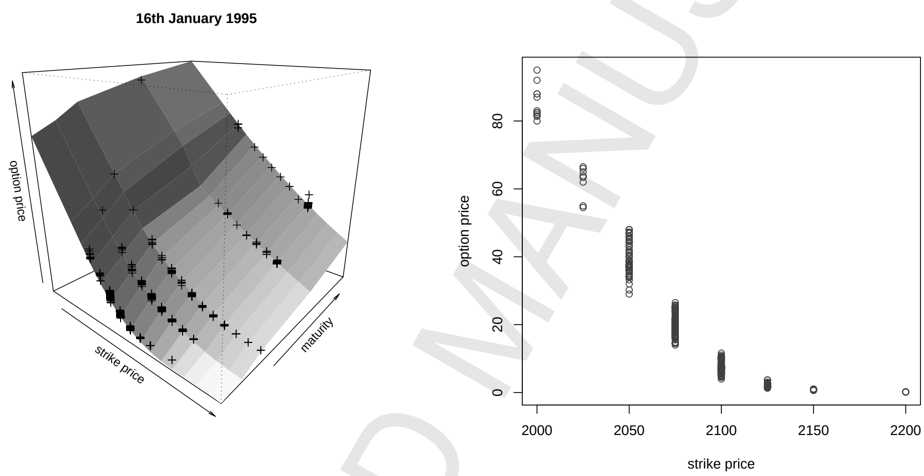


Fig. 3. Option prices plotted against strike price and time to maturity with two-dimensional kernel regression surface (left) in January 1995 and the ensemble of the call option prices with shortest time to expiry against strike price (right) on 16^{th} January 1995. SFB and CASE data base: sfb649.wiwi.hu-berlin.de.

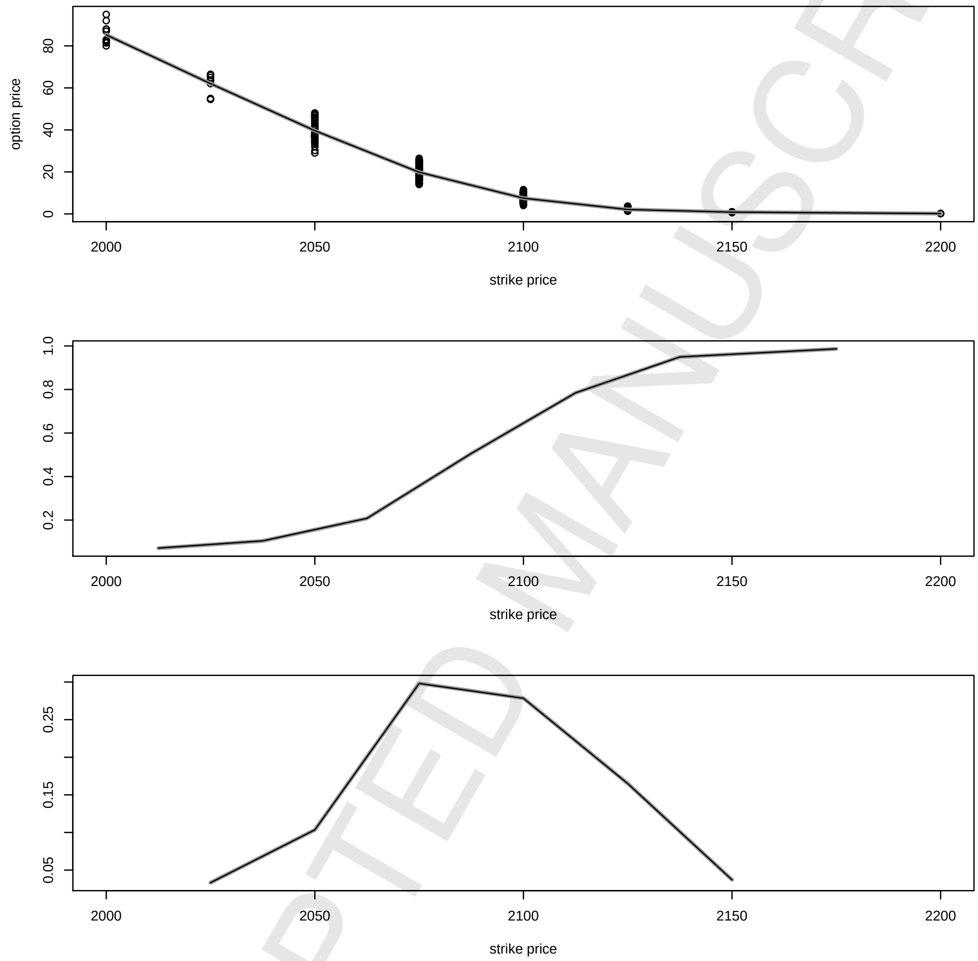


Fig. 4. On 16th January 1995, the unconstrained estimate satisfies the constraints. Hence, it is equal to the constrained estimate. The top panel shows the original data with the fitted call pricing functions. The second and the third panel show the estimates of the first and second derivative, respectively.

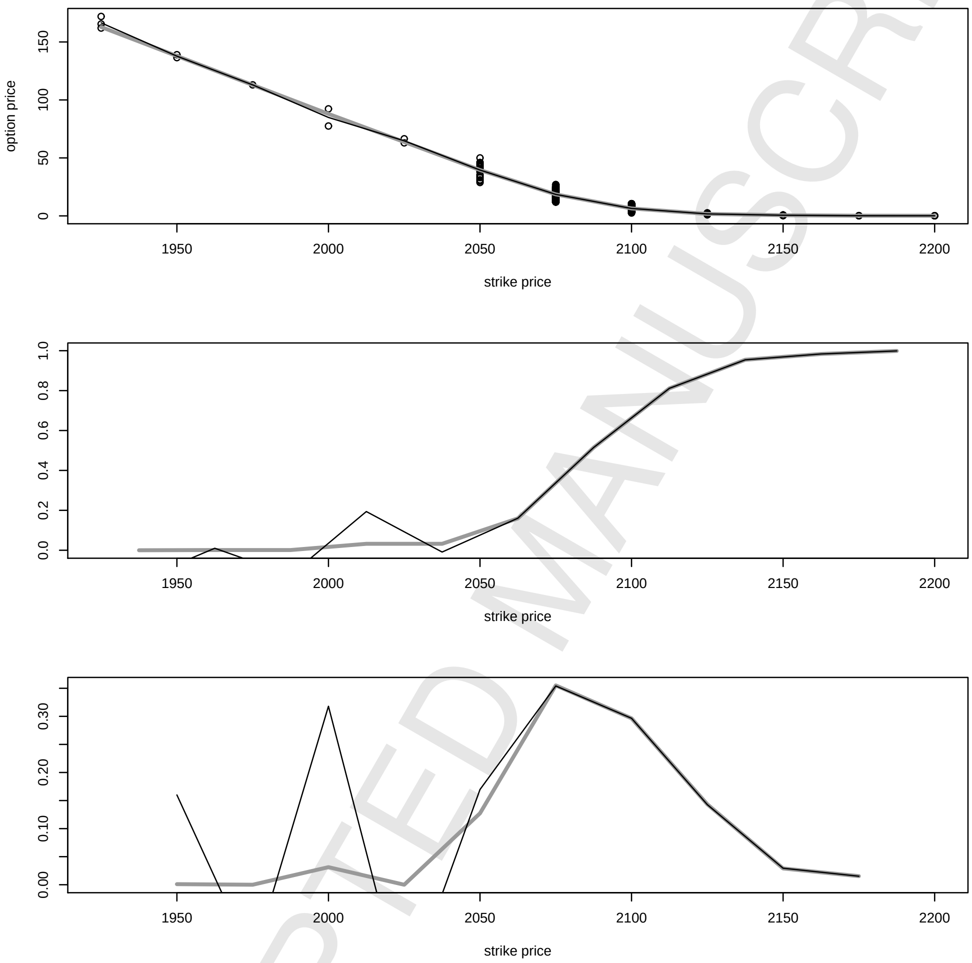


Fig. 5. On 17th January 1995, the unconstrained estimate, displayed using the thin line, does not satisfy the constraints. The top panel shows the original data with the two fitted call pricing functions. The estimates of the first derivative in the second panel look rather different. The constrained estimate of the second derivative in the bottom panel is clearly much more stable than the unconstrained estimate.

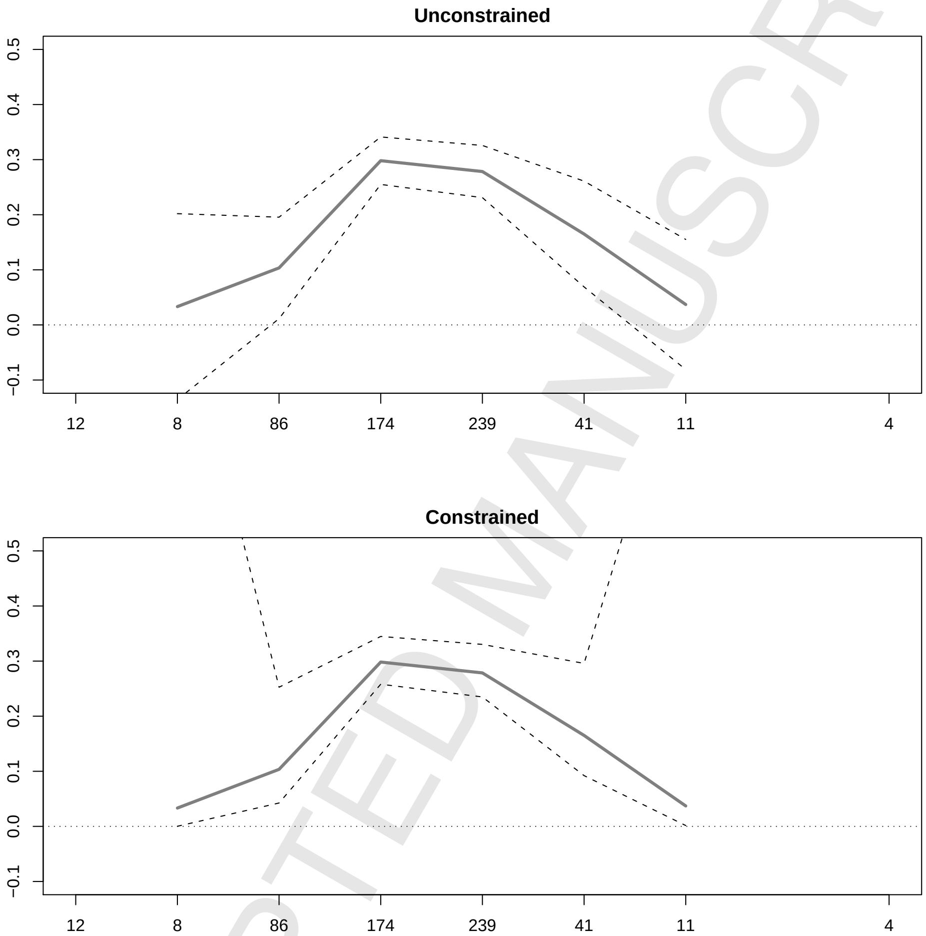


Fig. 6. The unconstrained and constrained confidence intervals for SPD on 16th January 1995. The description on the x-axis shows the number of observations in each point.

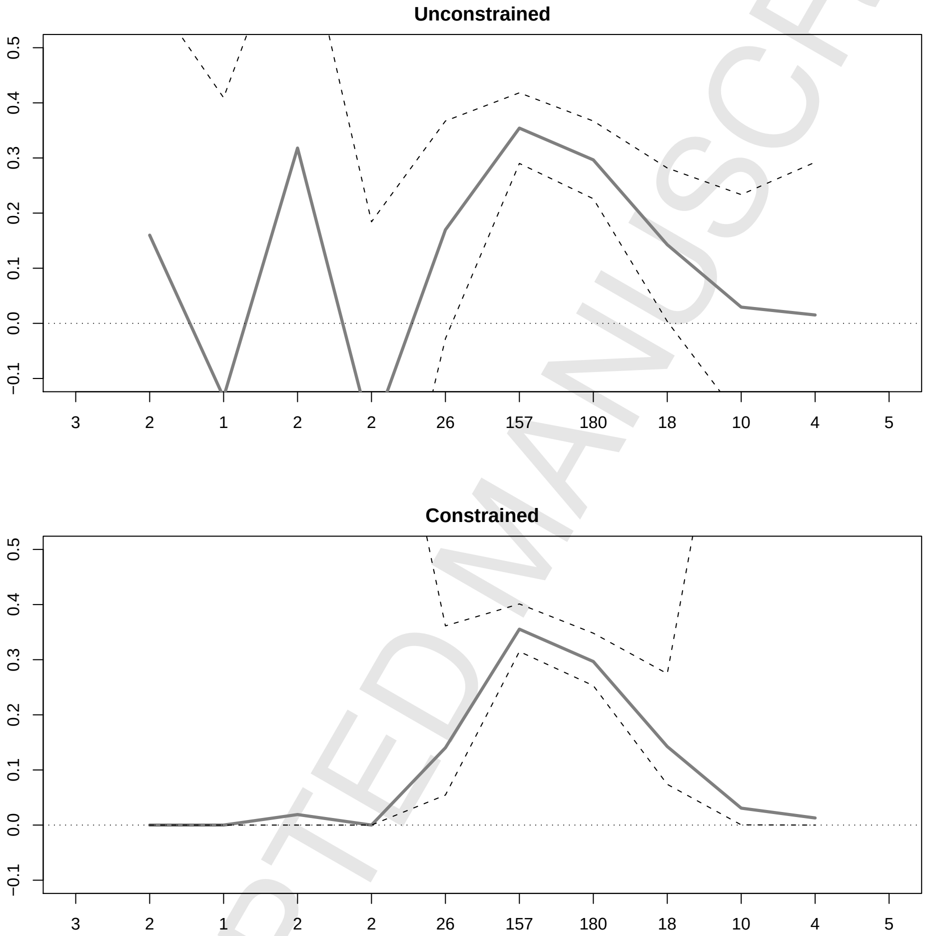


Fig. 7. Confidence intervals for SPD on 17th January 1995. The description on the x-axis shows the number of observations in each point.

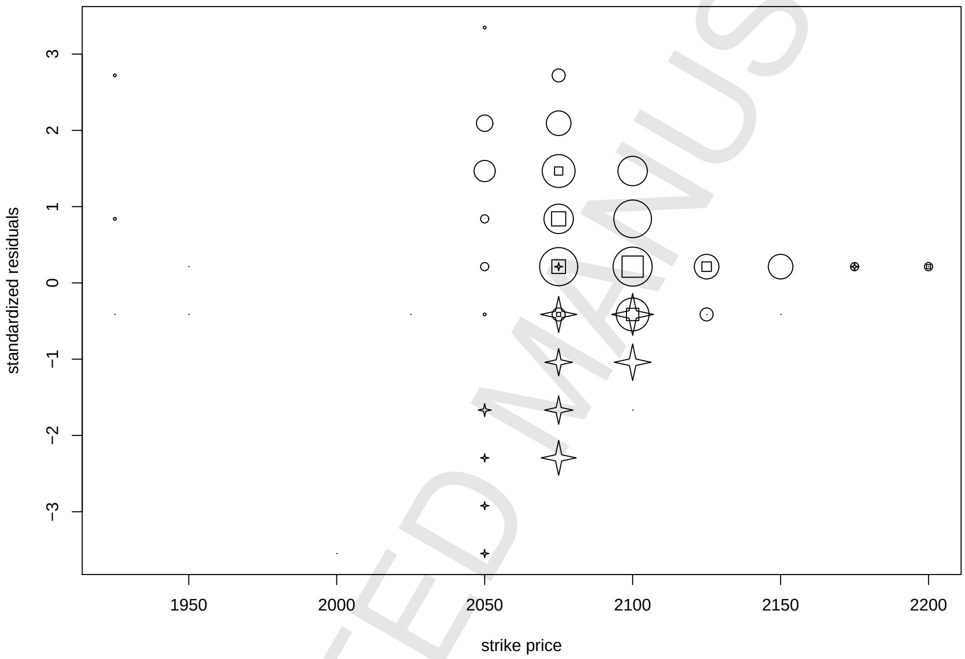


Fig. 8. The time dependency and the heteroscedasticity of the residuals during one day. The circle, square and the star denote the trades carried out in the morning, midday and in the afternoon, respectively. Size of the symbols denotes number of residuals.

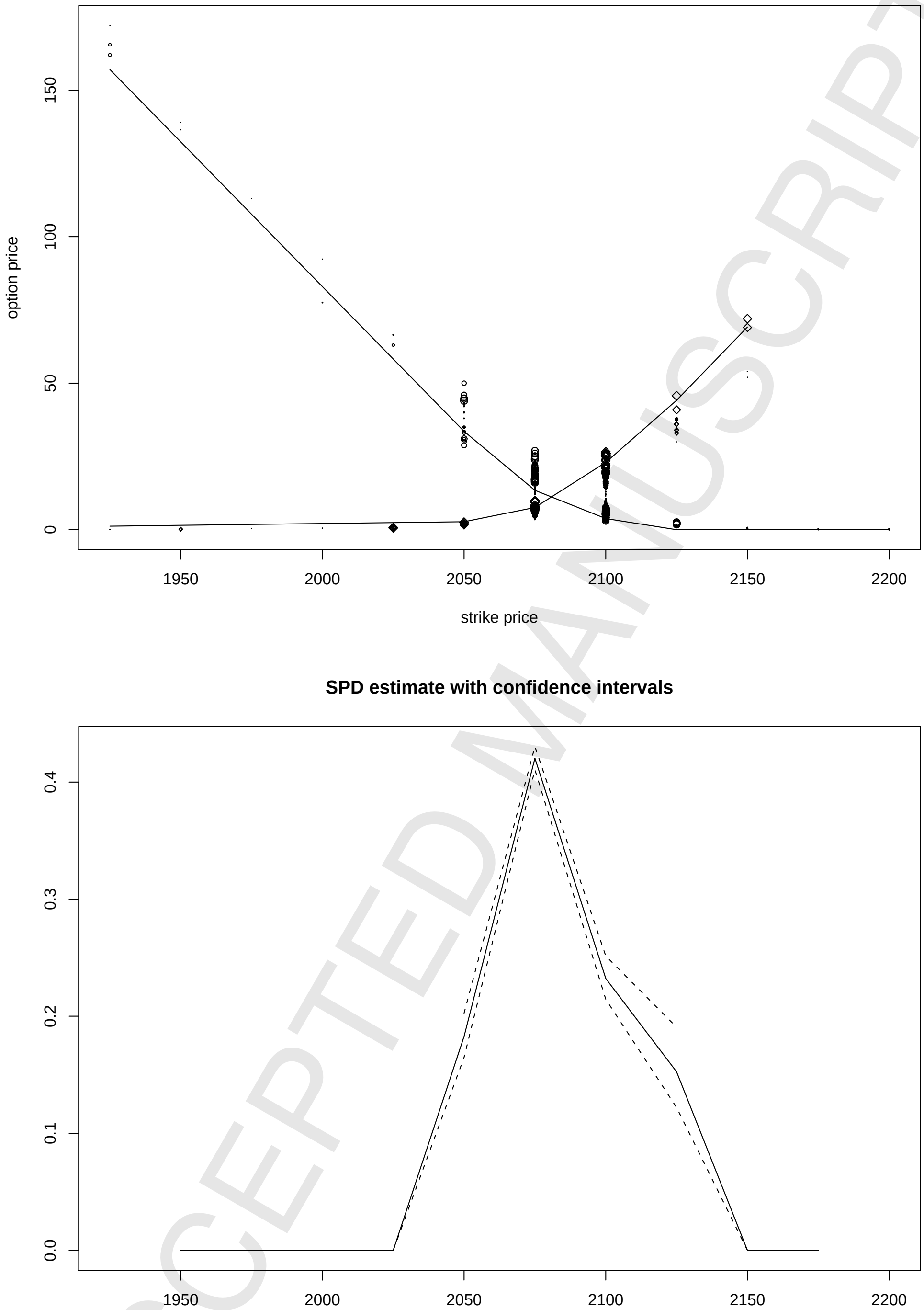
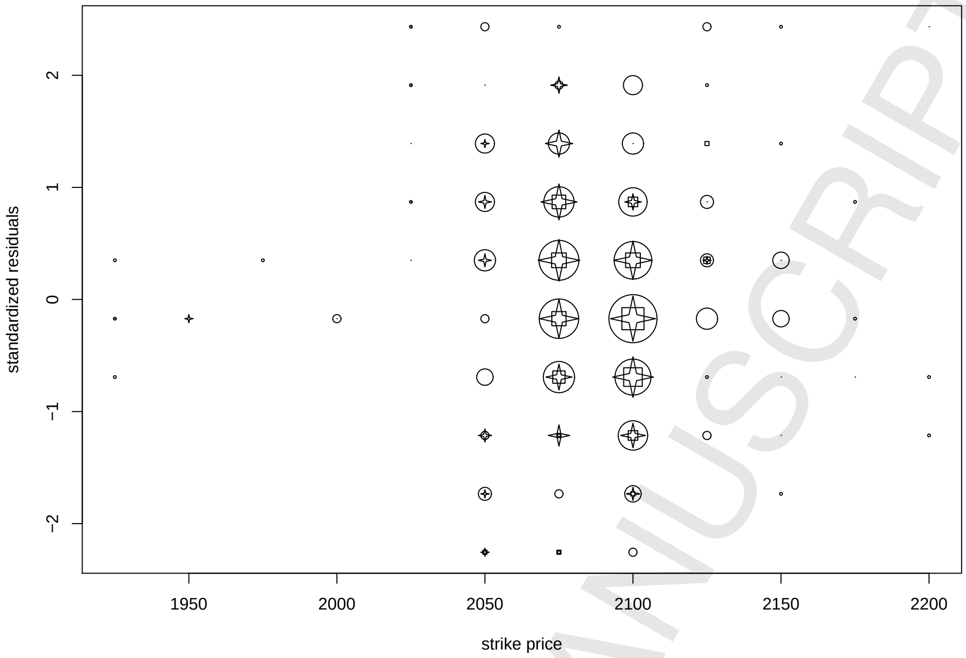


Fig. 9. Estimate using the covariance structure (28) on 17th January. The upper plot shows the observed option prices and the constrained estimate. The size of the plotting symbols corresponds to the weight of the observations. The lower plot shows estimated SPD with confidence intervals.



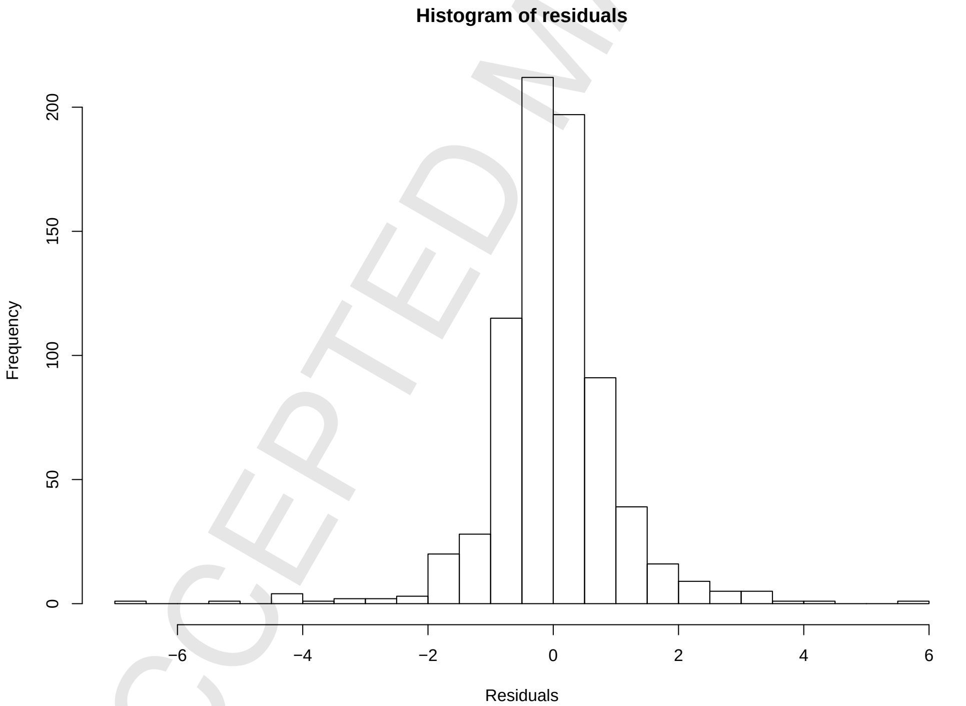


Fig. 10. The development of the standardized residuals resulting from model with the covariance structure (28) on 17th January during the day where circles, squares, and stars denote the residuals from morning, midday, and afternoon and a histogram of the standardized residuals.

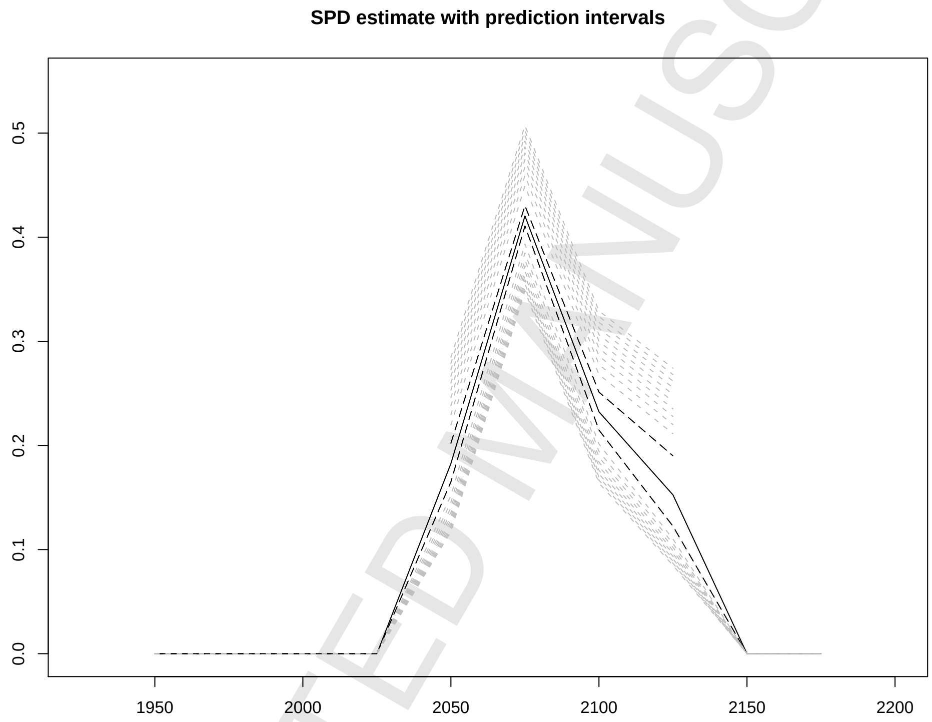


Fig. 11. SPD estimate on 17th January 1995 with prediction intervals for next 5 hours calculated for every 30 minutes.

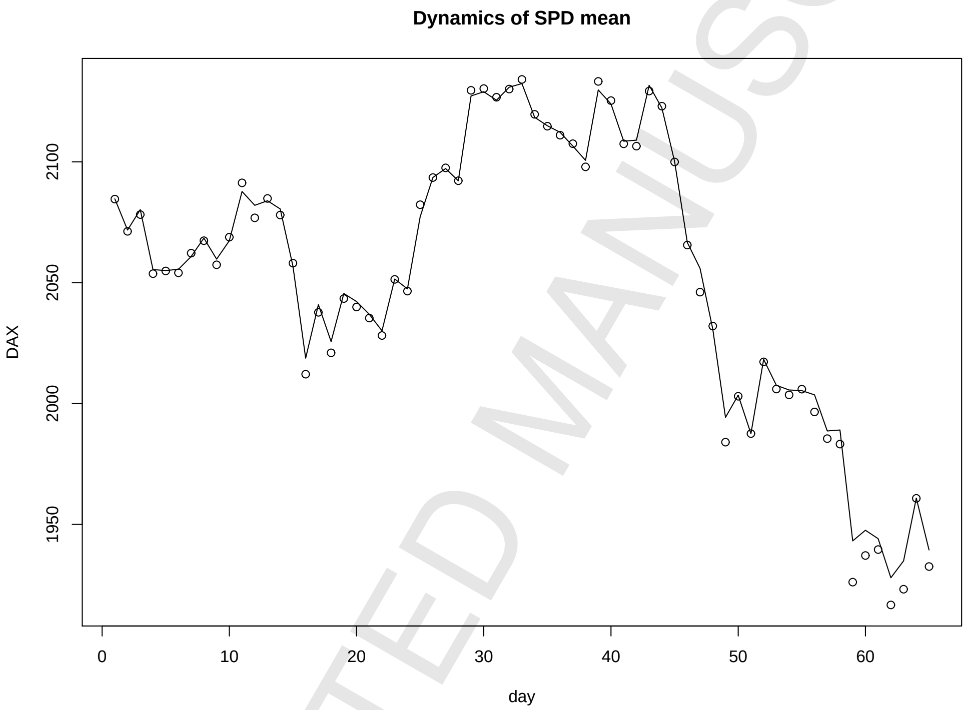


Fig. 12. Daily development of the expected value of the uncorrected SPD from January till March 1995. The circles denote the corresponding closing value of DAX.

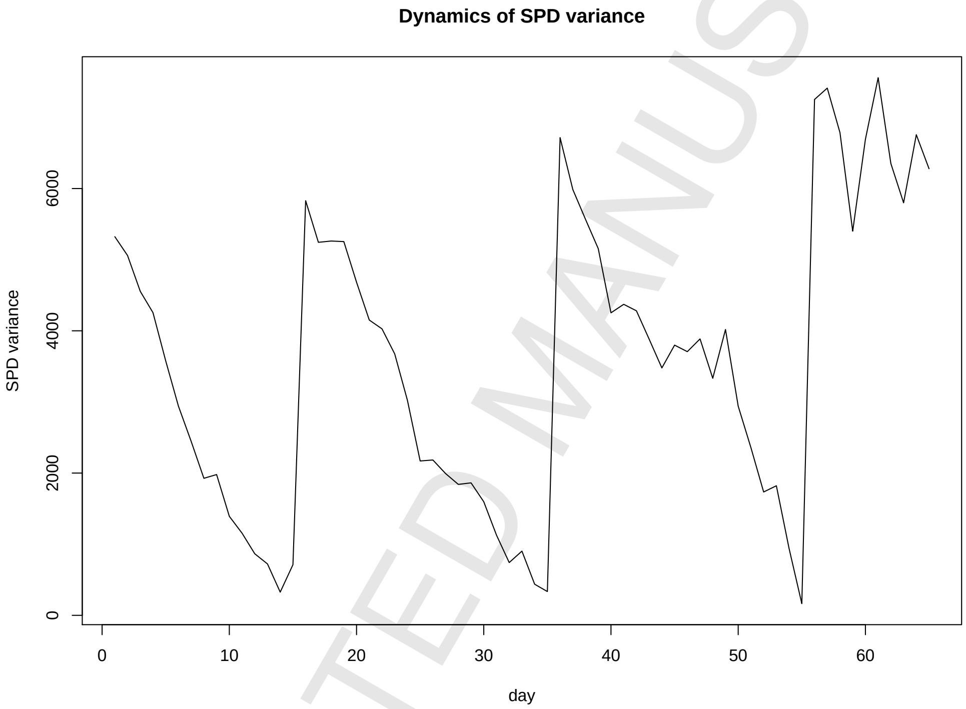


Fig. 13. Daily development of the SPD variance from January till March 1995.

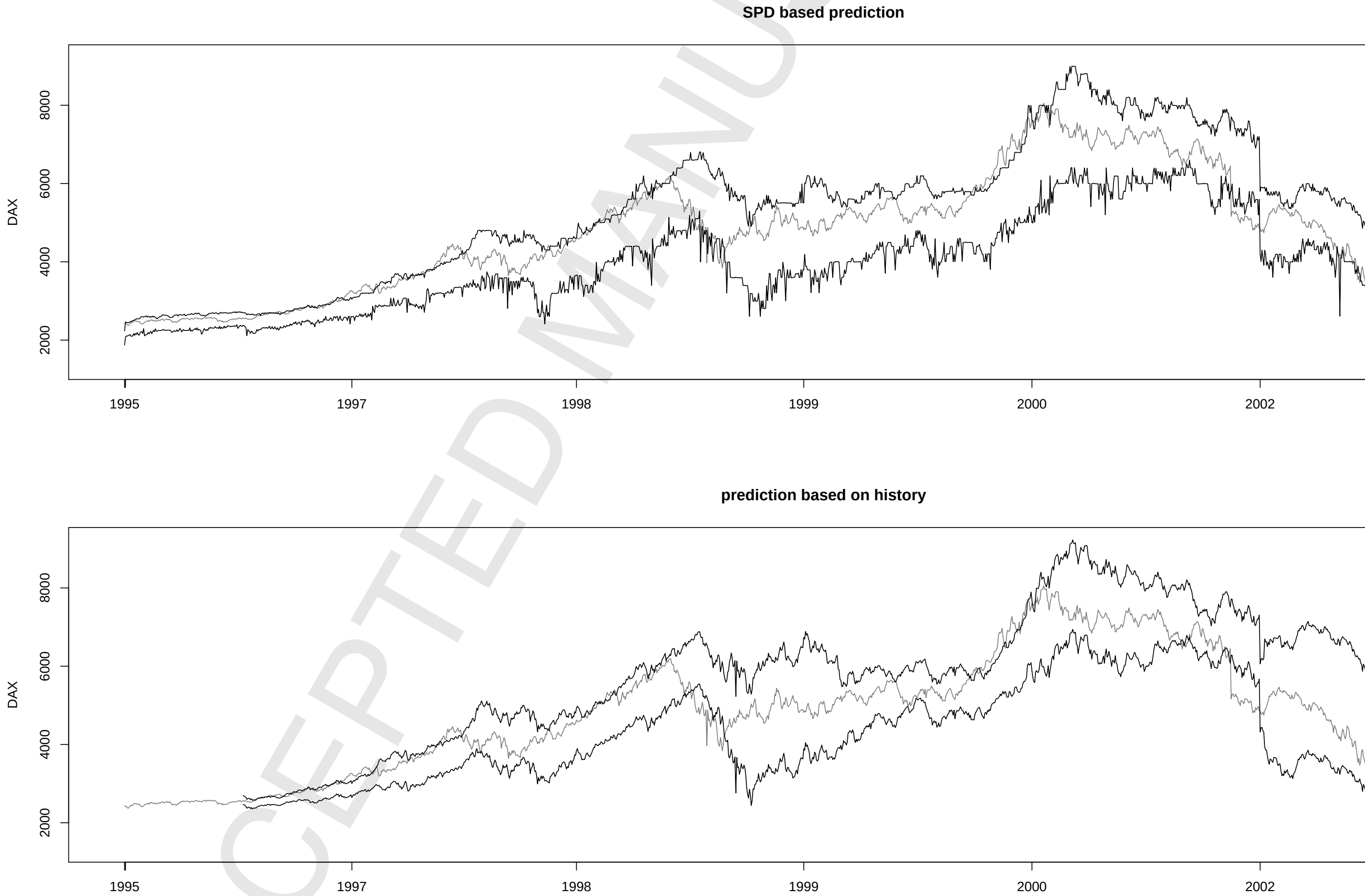


Fig. 14. Prediction intervals for DAX based on SPDs and historical simulation from January 1995 till

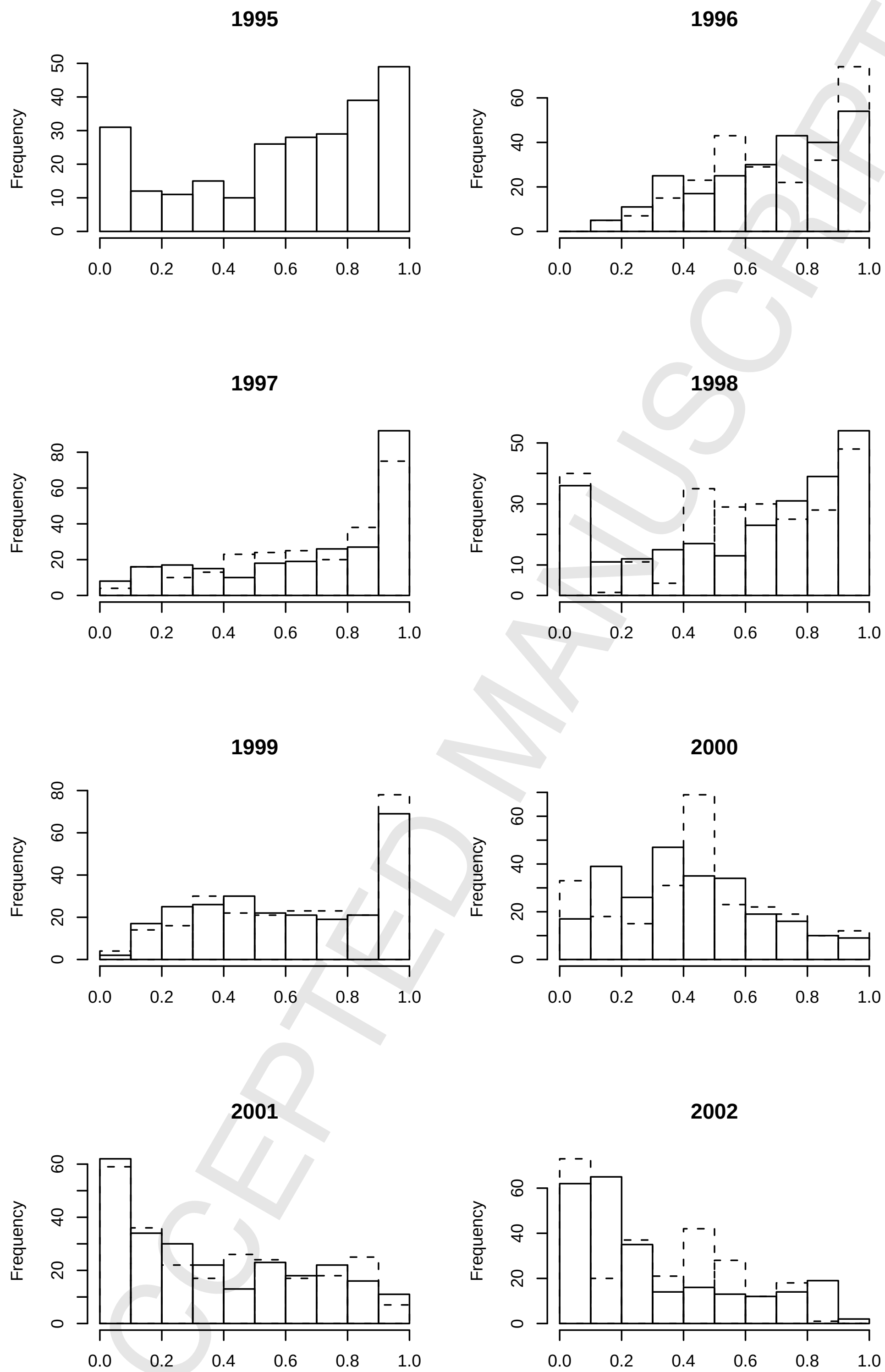


Fig. 15. Histograms for the SPDs (full line) and historical simulation (dashed line).

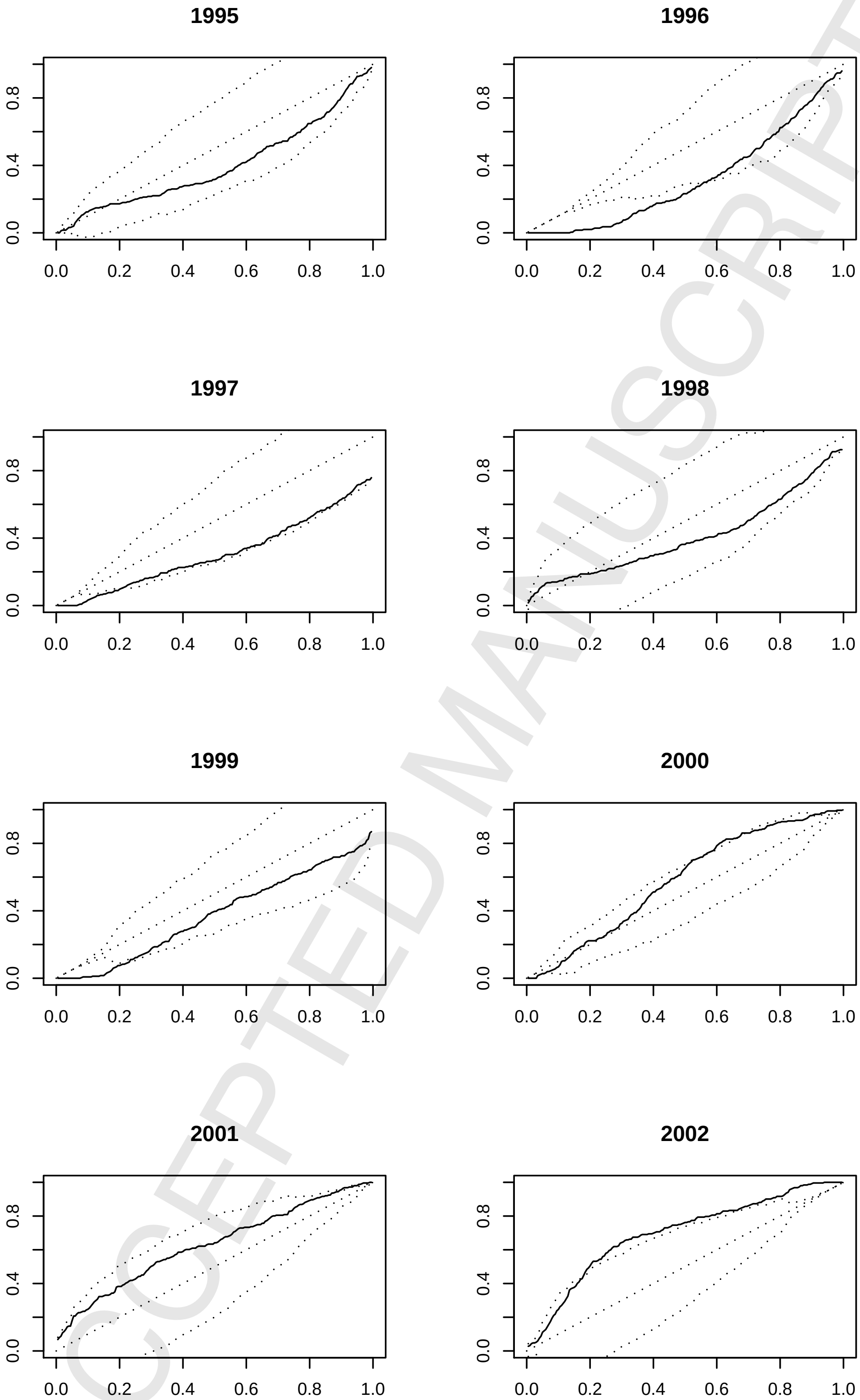


Fig. 16. Integral transformation for estimated SPDs.

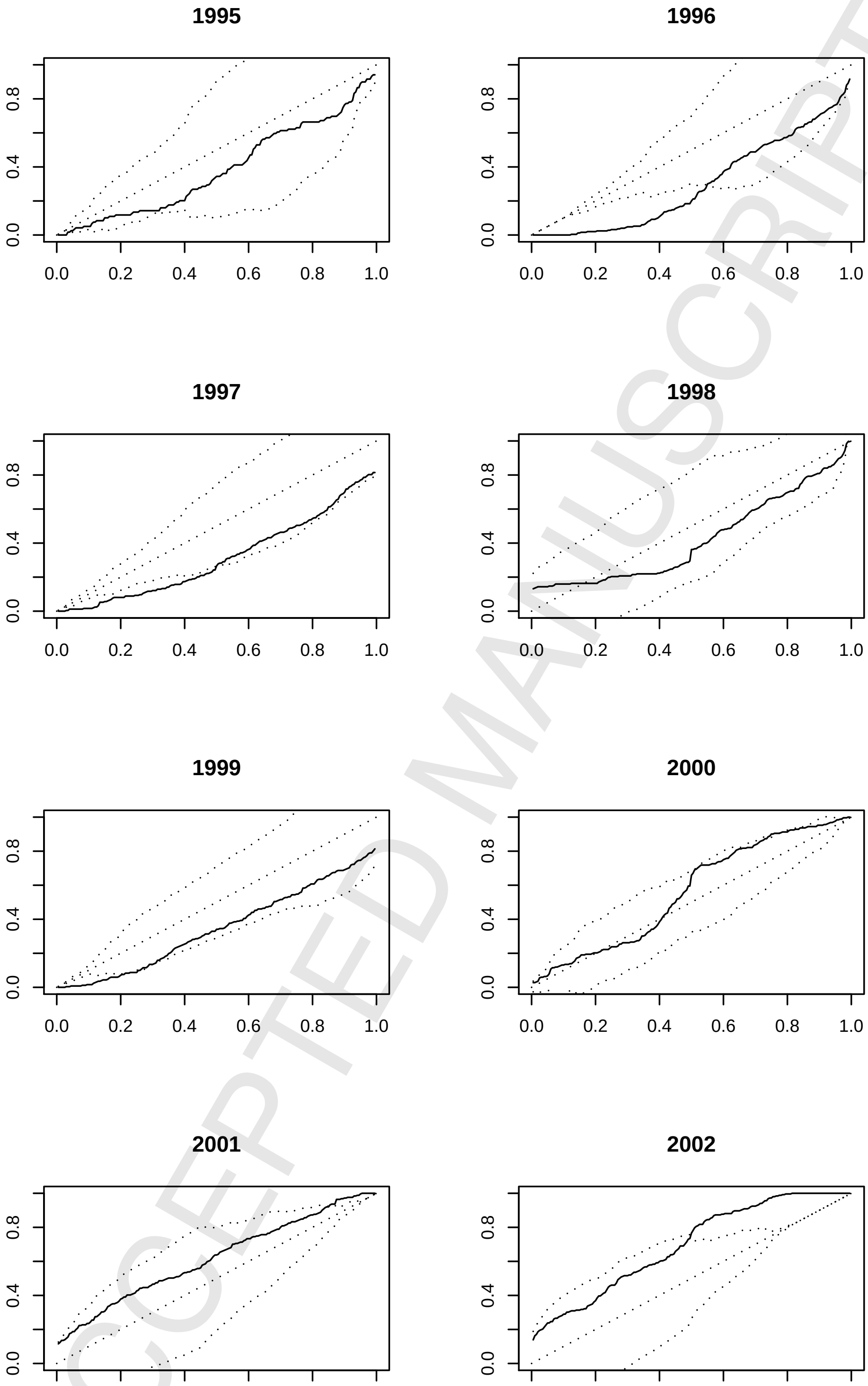


Fig. 17. Integral transformation for historical simulation.